

GenCore version 5.1.4\_P5\_4578  
 copyright (c) 1993 - 2003 Compugen Ltd.

score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

## SUMMARIES

OM nucleic - nucleic search, using sw model.					
Run on:	May 10, 2003, 15:03:13 ; Search time 3307 Seconds (without alignments)				
Perfect score:	9847.608 Million cell updates/sec				
Sequence:	1 atgatataaccaggctccaat.....ttcagcagggtgaaaaataa 1119				
Scoring table:	IDENTITY_NUC Gapop 16.0 , Gapext 1.0				
Title:	US-09-921-992-3				
Searched:	2054640 seqs, 14551402878 residues				
Total number of hits satisfying chosen parameters:	4109280				
Minimum DB seq length:	0				
Maximum DB seq length:	200000000				
Post-processing:	Minimum Match 0% Maximum Match 100% Listing first 45 summaries				
Database :					
1: GenEmbl:*					
2: gb_ba:*					
3: gb_htg:*					
4: gb_in:*					
5: gb_om:*					
6: gb_ov:*					
7: gb_pat:*					
8: .gb_ph:*					
9: .gb_p1:*					
10: .gb_pr:*					
11: .gb_sts:*					
12: .gb_sy:*					
13: .gb_un:*					
14: .gb_vl:*					
15: .em_ba:*					
16: .em_fun:*					
17: .em_hum:*					
18: .em_in:*					
19: .em_mu:*					
20: .em_on:*					
21: .em_or:*					
22: .em_ov:*					
23: .em_pat:*					
24: .em_ph:*					
25: .em_pl:*					
26: .em_ro:*					
27: .em_sts:*					
28: .em_un:*					
29: .em_htg_hum:*					
30: .em_htg_inv:*					
31: .em_htg_other:*					
32: .em_htg_mus:*					
33: .em_htg_pin:*					
34: .em_htg_rnd:*					
35: .em_htg_men:*					
36: .em_htg_vrt:*					
37: .em_sv:*					
38: .em_higo_hum:*					
39: .em_higo_mus:*					
40: .em_higo_other:*					
41: .em_higo_other:*					

Pred. No. 19 is the number of results predicted by chance to have a

BASE COUNT	272	a	279	c	317	9	251	t
Query Match	100.0%		Score 1119; DB 6;	Length 1119;				
Best Local Similarity	100.0%		Pred. No. 9.9e-271;					
Matches 1119; Conservative			0; Mismatches 0;					
Indels 0;			Gaps 0;					
QY	1	AATGCAATAACCAGCTCCATTCAACGTTAGAATCACACCTTACGTAGTGGAAATG	60					
Db	1	ATGCAATAACCAGCTCCATTCAACGTTAGAATCACACCTTACGTAGTGGAAATG	60					
QY	61	CGGATTGGCGATGGTCTGCATCGCGTACAGTCATGACCAATACGGTACAGAC	120					
Db	61	CGGATTGGCGATGGTCTGCATCGCGTACAGTCATGACCAATACGGTACAGAC	120					
QY	121	GTGGAAGAACGGTCAATGAAATCAAGGCGTGGAAACGGCTGGCTGATATCCTCGT	180					
Db	121	GTGGAAGAACGGTCAATGAAATCAAGGCGTGGAAACGGCTGGCTGATATCCTCGT	180					
QY	181	GTATCCGCTACGAGCTGGCTGACACACTCGACTATCGCATGGCTGAAGTAGCGAATAC	240					
Db	181	GTATCCGCTACGAGCTGGCTGACACACTCGACTATCGCATGGCTGAAGTAGCGAATAC	240					
QY	241	GGCGGTTGATGTCCTCGTAACTACCTTGCAATTCGCAATGGCTGAAGTAGCGAATAC	300					
Db	241	GGCGGTTGATGTCCTCGTAACTACCTTGCAATTCGCAATGGCTGAAGTAGCGAATAC	300					
QY	301	GGCGGTTGATGTCCTCGTAACTACCTTGCAATTCGCAATGGCTGAAGTAGCGAATAC	360					
Db	301	GGCGGTTGATGTCCTCGTAACTACCTTGCAATTCGCAATGGCTGAAGTAGCGAATAC	360					
QY	361	GGGGTTGACTGGCGGAAACACATTCGATCCGTTAACACATTCGATCCGTTAACCGGATCG	420					
Db	361	GGGGTTGACTGGCGGAAACACATTCGATCCGTTAACACATTCGATCCGTTAACCGGATCG	420					
QY	421	CTGGAAAATGATCNGCAACAAACTATGGCGAACCGACGCC3AGGGTTGCTGAATCT	480					
Db	421	CTGGAAAATGATCNGCAACAAACTATGGCGAACCGACGCC3AGGGTTGCTGAATCT	480					
QY	481	GGCTCTGCGTCAATGAAATCAACGTTGAAATCACACCTTACGTGGAAATG	540					
Db	481	GGCTCTGCGTCAATGAAATCAACGTTGAAATCACACCTTACGTGGAAATG	540					
QY	541	AAGGCTCTGACCTCTCTGGCTGTAAGCTTATCCCTCTGGAAAACAGATCGAT	600					
Db	541	AAGGCTCTGACCTCTCTGGCTGTAAGCTTATCCCTCTGGAAAACAGATCGAT	600					
QY	601	CAGCGGTGCACTGGGCAACGCCGAAGCCGGTGGCGCAGAGCGGGCAATTAATCC	660					
Db	601	CAGCGGTGCACTGGGCAACGCCGAAGCCGGTGGCGCAGAGCGGGCAATTAATCC	660					
QY	660	CAACCGTTGCACTGGGCAACGCCGAAGCCGGTGGCGCAGAGCGGGCAATTAATCC	660					
Db	660	CAACCGTTGCACTGGGCAACGCCGAAGCCGGTGGCGCAGAGCGGGCAATTAATCC	660					
QY	661	GGCATTTGGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	720					
Db	661	GGCATTTGGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	720					
QY	721	GGGGCGATGGCTGAGAGCTCAACTGCTGGCTGAGAGCTCAACTGCTGGCTG	780					
Db	721	GGGGCGATGGCTGAGAGCTCAACTGCTGGCTGAGAGCTCAACTGCTGGCTG	780					
QY	781	CGTCGCGAGGGATCACTGATCAGCTGGCTGCGGACGCTGTCGCTGCTGCTG	840					
Db	781	CGTCGCGAGGGATCACTGATCAGCTGGCTGCGGACGCTGTCGCTGCTGCTG	840					
QY	840	CGTCGCGAGGGATCACTGATCAGCTGGCTGCGGACGCTGTCGCTGCTGCTG	840					
Db	840	CGTCGCGAGGGATCACTGATCAGCTGGCTGCGGACGCTGTCGCTGCTGCTG	840					
QY	841	ATCGTGTACGGTTAACGGCGTGGGAAACGGCTGGAGATACTCATCGATGGACGT	900					
Db	841	ATCGTGTACGGTTAACGGCGTGGGAAACGGCTGGAGATACTCATCGATGGACGT	900					
QY	900	TGCAATGTTGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	960					
Db	900	TGCAATGTTGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	960					
QY	960	ACCGGGGGAAACAGAAAGGGCCTCTATGAGATGGCGTGGCAAGAACGCTCTGGAC	1020					
Db	960	ACCGGGGGAAACAGAAAGGGCCTCTATGAGATGGCGTGGCAAGAACGCTCTGGAC	1020					
QY	1021	AACACCATATGATCGGCCACAGCTGGAAAGCAGCATCGCTGAC	1080					
Db	1021	AACACCATATGATCGGCCACAGCTGGAAAGCAGCATCGCTGAC	1080					
QY	1081	GAAGCGGTCAAATGAGCTGAGCTGAGTGGAAATAA	1119					
Db	1081	GAAGCGGTCAAATGAGCTGAGCTGAGTGGAAATAA	1119					
RESULT 2								
AX038912								
LOCUS								
DEFINITION								
Sequence 29 from Patent WO0061793.								
ACCESSION								
AX038912								
VERSION								
KEYWORDS								
SOURCE								
Escherichia coli.								
ORGANISM								
Bacterium; Proteobacteria; gamma subdivision; Enterobacteriaceae;								
REFERENCE								
(bases 1 to 1119)								
AUTHORS								
Loferer, H. and Jacobi, A.								
JOURNAL								
Novel method for identifying antibacterial compounds								
FEATURES								
Source /organism="Escherichia coli"								
BASE COUNT	272	a	279	c	317	9	251	t
ORIGIN								
Query Match	100.0%		Score 1119; DB 6;	Length 1119;				
Best Local Similarity	100.0%		Pred. No. 9.9e-271;					
Matches 1119; Conservative			0; Mismatches 0;					
Indels 0;			Gaps 0;					
QY	1	ATGCAATAACCAGCTCCATTCAACGTTAGAATCACACCTTACGTAGTGGAAATG	60					
Db	1	ATGCAATAACCAGCTCCATTCAACGTTAGAATCACACCTTACGTAGTGGAAATG	60					
QY	61	CGGATTGGCGATGGTCTGCATCGCGTACAGTCATGACCAATACGGTACAGAC	120					
Db	61	CGGATTGGCGATGGTCTGCATCGCGTACAGTCATGACCAATACGGTACAGAC	120					
QY	121	GTGGAAGAACGGTCAATGAAATCAACGTTAGAATCACACCTTACGTAGTGGAAATG	180					
Db	121	GTGGAAGAACGGTCAATGAAATCAACGTTAGAATCACACCTTACGTAGTGGAAATG	180					
QY	181	GTATCCGCTACGAGCTGGCTGACACACTCGACTATCGCATGGCTGAAGTAGCGAATAC	240					
Db	181	GTATCCGCTACGAGCTGGCTGACACACTCGACTATCGCATGGCTGAAGTAGCGAATAC	240					
QY	240	GGCGGTTGATGTCCTCGTAACTACCTTGCAATTCGCAATGGCTGAAGTAGCGAATAC	300					
Db	240	GGCGGTTGATGTCCTCGTAACTACCTTGCAATTCGCAATGGCTGAAGTAGCGAATAC	300					
QY	300	GGCGGTTGATGTCCTCGTAACTACCTTGCAATTCGCAATGGCTGAAGTAGCGAATAC	360					
Db	300	GGCGGTTGATGTCCTCGTAACTACCTTGCAATTCGCAATGGCTGAAGTAGCGAATAC	360					
QY	360	GGGGTTGACTGGCGGAAACACATTCGATCCGTTAACACATTCGATCCGTTAACCGGATCG	420					
Db	360	GGGGTTGACTGGCGGAAACACATTCGATCCGTTAACACATTCGATCCGTTAACCGGATCG	420					
QY	420	CTGGAAAATGATCNGCAACAAACTATGGCGAACCGACGCC3AGGGTTGCTGAATCT	480					
Db	420	CTGGAAAATGATCNGCAACAAACTATGGCGAACCGACGCC3AGGGTTGCTGAATCT	480					
QY	480	GGCTCTGCGTCAATGAAATCAACGTTGAAATCACACCTTACGTGGAAATG	540					
Db	480	GGCTCTGCGTCAATGAAATCAACGTTGAAATCACACCTTACGTGGAAATG	540					
QY	540	AAGGCTCTGACCTCTCTGGCTGTAAGCTTATCCCTCTGGAAAACAGATCGAT	600					
Db	540	AAGGCTCTGACCTCTCTGGCTGTAAGCTTATCCCTCTGGAAAACAGATCGAT	600					
QY	600	CAACCGTTGCACTGGGCAACGCCGAAGCCGGTGGCGCAGAGCGGGCAATTAATCC	660					
Db	600	CAACCGTTGCACTGGGCAACGCCGAAGCCGGTGGCGCAGAGCGGGCAATTAATCC	660					
QY	660	GGCATTTGGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	720					
Db	660	GGCATTTGGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	720					
QY	720	GGGGCGATGGCTGAGAGCTCAACTGCTGGCTGAGAGCTCAACTGCTGGCTG	780					
Db	720	GGGGCGATGGCTGAGAGCTCAACTGCTGGCTGAGAGCTCAACTGCTGGCTG	780					
QY	780	CGTCGCGAGGGATCACTGATCAGCTGGCTGCGGACGCTGTCGCTGCTGCTG	840					
Db	780	CGTCGCGAGGGATCACTGATCAGCTGGCTGCGGACGCTGTCGCTGCTGCTG	840					
QY	840	ATCGTGTACGGTTAACGGCGTGGGAAACGGCTGGAGATACTCATCGATGGACGT	900					
Db	840	ATCGTGTACGGTTAACGGCGTGGGAAACGGCTGGAGATACTCATCGATGGACGT	900					
QY	900	TGAGATGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	960					
Db	900	TGAGATGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	960					
QY	960	ACCGGGGGAAACAGAAAGGGCCTCTATGAGATGGCGTGGCAAGAACGCTCTGGAC	1020					
Db	960	ACCGGGGGAAACAGAAAGGGCCTCTATGAGATGGCGTGGCAAGAACGCTCTGGAC	1020					
QY	1021	AACACCATATGATCGGCCACAGCTGGAAAGCAGCATCGCTGAC	1080					
Db	1021	AACACCATATGATCGGCCACAGCTGGAAAGCAGCATCGCTGAC	1080					
QY	1081	GAAGCGGTCAAATGAGCTGAGCTGAGTGGAAATAA	1119					
Db	1081	GAAGCGGTCAAATGAGCTGAGCTGAGTGGAAATAA	1119					
QY	1119	GGCGGTTGACTGGCGGAAACACATTCGATCCGTTAACACATTCGATCCGTTAACCGGATCG	120					
Db	1119	GGCGGTTGACTGGCGGAAACACATTCGATCCGTTAACACATTCGATCCGTTAACCGGATCG	120					
QY	120	GTGGAAGAACGGTCAATGAAATCAACGTTGAAATCACACCTTACGTGGAAATG	180					
Db	120	GTGGAAGAACGGTCAATGAAATCAACGTTGAAATCACACCTTACGTGGAAATG	180					
QY	180	GGCGGTTGACTGGCTGAAACACATTCGATCCGTTAACACATTCGATCCGTTAACCGGATCG	240					
Db	180	GGCGGTTGACTGGCTGAAACACATTCGATCCGTTAACACATTCGATCCGTTAACCGGATCG	240					
QY	240	GGCGGTTGACTGGCTGAAACACATTCGATCCGTTAACACATTCGATCCGTTAACCGGATCG	300					
Db	240	GGCGGTTGACTGGCTGAAACACATTCGATCCGTTAACACATTCGATCCGTTAACCGGATCG	300					
QY	300	GGCGGTTGACTGGCTGAAACACATTCGATCCGTTAACACATTCGATCCGTTAACCGGATCG	360					
Db	300	GGCGGTTGACTGGCTGAAACACATTCGATCCGTTAACACATTCGATCCGTTAACCGGATCG	360					
QY	360	GGGGTTGACTGGCGGAAACACATTCGATCCGTTAACACATTCGATCCGTTAACCGGATCG	420					
Db	360	GGGGTTGACTGGCGGAAACACATTCGATCCGTTAACACATTCGATCCGTTAACCGGATCG	420					
QY	420	GTGGAAGAACGGTCAATGAGCTGAGCTGAGTGGAAATAA	480					
Db	420	GTGGAAGAACGGTCAATGAGCTGAGCTGAGTGGAAATAA	480					
QY	480	GGCGGTTGACTGGCTGAAACACATTCGATCCGTTAACACATTCGATCCGTTAACCGGATCG	540					
Db	480	GGCGGTTGACTGGCTGAAACACATTCGATCCGTTAACACATTCGATCCGTTAACCGGATCG	540					





JOURNAL Submitted (17-FEB-1992) J. Parker, Southern Illinois Univ., Dept of Microbiology, S I U, Carbondale IL 62901, USA  
 REFERENCE 2 (bases 1 to 1697)  
 AUTHORS Baker,J., Franklin,D.B. and Parker,J.  
 TITLE Sequence and characterization of the *gcpE* gene of *Escherichia coli*.  
 JOURNAL FEMS Microbiol. Lett. 94, 175-180 (1992)  
 COMMENT See also J01629 & M11843.  
 FEATURES Location/Qualifiers  
 SOURCE . 1697  
*/organism="Escherichia coli"*  
*/strain="K12"*  
*/db\_xref="txon:562"*  
*/sub\_clone="PSE411"*  
*/cell\_line="JF298"*  
*/clone\_libr="lambda transducing phage dguAB-2"*  
*<1..508*  
*note="orf"*  
*/codon\_start=2*  
*/transl\_table=11*  
*/protein\_id="CAA45782.1"*  
*/db\_xref="GI:41541"*  
*/db\_xref="SWISS-PROT:P27434"*  
*/translation="TTDPAITSPPPAVDTATNTOPPAVAPAPA PAVDPOOQNAVSP.*  
*SQANTISDAMPAPTAATTGDAALPLPTDAGWTPVADPNVAMNPFADCNIEYTDAT*  
*GKLFSGMQRKKDGNIULIQPAVRLKIGAPAVQIQTQKGPKPDLSRFTRNQVARLIL*  
*NAEQSPAQ"*  
 gene 535 .. 1653  
*/gene="gcpE"*  
*/35.. 1653*  
*/codon\_start=1*  
*/transl\_table=11*  
*/protein\_id="CAA45783.1"*  
*/db\_xref="GI:41542"*  
*/db\_xref="SWISS-PROT:P27433"*  
*/translation="MINAPQPRRSTIYGVNUVPIGDAPPVIAQVSMTNTRDVEAT*  
*VNQIKALERVGADIVRVSPVMPMDAKELKLKQVNIVPLVADHDFYRALKVAEYGV*  
*DCLRINPQNEERTRMVVCARDINLPVTRIVNGSLERKLDQERIGEPTPQLVES*  
*AMHRVHDLDNFEDOFEVVUSCDAVINLPIVYLNQAKIDQPHLIGITRAGGANSQAV*  
*KSAHGIGLILSSEIGDILRVLSSADDEVEEIKVGFDLKSIRSRGINTIACPCSRQ*  
*EFVIGTGNNALQSRLEDITMDVSDISLGIVCNGPGRALVSLTGIVGNKRKSLVYEDGV*  
*RKORLDDNMIDOLEAATRAKASOLBARRDVOQEK"*  
 BASE COUNT 417 a 446 c 469 g 365 t  
 ORIGIN 446 c 469 g 365 t

Query Match 100.0% Score 1119; DB 1; Length 1697;  
 Best Local Similarity 100.0%; Pred. NO. 1e 270; Matches 1119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 ATGCATAAACCGGGCTCAATTCAAGCTAGAAATCACAGCTATTAGCTGGATGT 60  
 Db 355 ATGCATAACCGGGCTCAATTCAAGCTAGAAATCACAGCTATTAGCTGGATGT 594  
 Qy 61 CGGATTCGGCGATGTCCTCCATCGCGGTACAGTCCATGACAAATACCGTAGACAGAC 120  
 Db 595 CGGATTCGGCGATGTCCTCCATCGCGGTACAGTCCATGACAAATACCGTAGACAGAC 654  
 Qy 121 GTGCGAAGCAAGGTCAATCAAACTCAAGGCCGTGGACGCTGGCTGATATGTCGT 180  
 Db 655 GTGCGAAGCAAGGTCAATCAAACTCAAGGCCGTGGACGCTGGCTGATATGTCGT 714  
 Qy 181 GTATCCGTACGACGAGTGAAGGGCGAGAACGTTCAACCTCATCAACAGCAGTTAAC 240  
 Db 715 GTATCCGTACGACGAGTGAAGGGCGAGAACGTTCAACCTCATCAACAGCAGTTAAC 774  
 Qy 241 GTGCCGCTGTGGCTGACATCCACTCTGACATATCGCATGGCTGAAAGTAGCGGAATAAC 300  
 Db 775 GTGCCGCTGTGGCTGACATCCACTCTGACATATCGCATGGCTGAAAGTAGCGGAATAAC 834  
 Qy 301 GGCGTCGATGTCGCGTATTAACCTCGCAATATCGTAGAAGAGCGTATTGCGATG 360  
 Db 835 GGCGTCGATGTCGCGTATTAACCTCGCAATATCGTAGAAGAGCGTATTGCGATG 894

RESULT 6  
 AE000318/c  
 LOCUS AE000318  
 DEFINITION Escherichia coli K12 MG1655 section 228 of 400 of the complete genome.  
 ACCESSION AE000318  
 VERSION AE000318.1 GI:1788862  
 KEYWORDS SOURCE Escherichia coli K12.  
 ORGANISM Escherichia coli K12.  
 AUTHORS Blattner,F.R., Blunkett,G., III, Bloch,C.A., Perera,N.T., Burland,V., Riley,M., Collado-Vides,J., Glasner,J.D., Rode,C.K., Mayhew,G.F., Gregor,J., Davis,N.W., Kirkpatrick,H.A., Goeden,M.A., Rose,D.J., Mau,B. and Shao,Y.  
 TITLE The complete genome sequence of *Escherichia coli* K-12  
 JOURNAL Science 277 (5351), 1453-1474 (1997)  
 MEDLINE 9742617

Qy 361 GTGGGTGACTGTGGCGCGGATAAACATCCGATCGTATGGGTTAACGGCGGATCG 420  
 Db 895 GTGGGTGACTGTGGCGCGGATAAACATCCGATCGTATGGGTTAACGGCGGATCG 954  
 Qy 421 CTGCGAAAAGATCTGCGAAGAAAGTAGTGGCAACCGACGCGCAGGGCTGCTGGATCT 480  
 Db 955 CTGGAAAAGATCTGCGAAGAAAGTAGTGGCAACCGACGCGCAGGGCTGCTGGATCT 1014  
 Qy 481 GCGATGGGACATGTTGATCATCPCGATGCGCAGACTTCGTCAGTCAGTCAGCTG 540  
 Db 1015 GCGATGGTCAATGTCATCTCGATGCGTCAAGTCAGTCAGTCAGCTG 1074  
 Qy 541 AAAGGCTCTGGCTTCTCCGCTGAGTCATATGTTCTGGCAACAGATGAT 600  
 Db 1075 AAAGGCTCTGGCTTCTCCGCTTCTGGCTTCTGTTGATGCTTGTGGCAAAAGAGTCGAT 1134  
 Qy 601 CAGCCCTGGTCACTGGGATCACCGAAGCCGGTGGGGCGACGGCTGGCGTATCGCTG 660  
 Db 1135 CAGCCCTGGTCACTGGGATCACCGAAGCCGGTGGGGCGACGGCTGGCGTATCGCTG 1194  
 Qy 661 GCGATGGTTAGGTCTGCTGCTGAGGGCATCGCGACACGGCTGGCGTATCGCTG 720  
 Db 1195 GCGATGGTTAGGTCTGCTGCTGCTGAGGGCATCGCGACACGGCTGGCGTATCGCTG 1254  
 Qy 721 GGGGCCATCGGGTCAAGAGATCAAGTAGGGTGTGATTTGAATGCGTGCATTC 780  
 Db 1255 GGGGCCATCGGGTCAAGAGATCAAGTAGGGTGTGATTTGAATGCGTGCATTC 1314  
 Qy 781 CGTCGGAGGGATCAACTATCGCTGCGGACGCTGTGCGTCAAGGATTTGATGTT 840  
 Db 1315 CGTCGGAGGGATCAACTATCGCTGCGGACGCTGTGCGTCAAGGATTTGATGTT 1374  
 Qy 841 ATCGGTAGGTTAACGCGCTGGAGCACGCGCTGGAGATATCATCACTCGATGACGTT 900  
 Db 1375 ATCGGTAGGTTAACGCGCTGGAGCACGCGCTGGAGATATCATCACTCGATGACGTT 1434  
 Qy 901 TCGATTACGGCTCGTGTGTAATGGCCAGGTGGCGCTGTTCTACACTCGCTG 960  
 Db 1435 TCGATTACGGCTCGTGTGTAATGGCCAGGTGGCGCTGTTCTACACTCGCTG 1494  
 Qy 961 ACCGGGGGACAAGAAAGANGGCCCTPATGAGATGGCGGCGCAAGACGCTGCGAC 1020  
 Db 1495 ACCGGGGGACAAGAAAGANGGCCCTPATGAGATGGCGGCGCAAGACGCTGCGAC 1554  
 Qy 1021 AACAAAGGATAGTGGACCGCTGGAGAACCGCATGGCAAGAGCTGAC 1080  
 Db 1555 AACAAAGGATAGTGGACCGCTGGAGAACGGCTCATGGCAAGAGCTGAC 1614  
 Qy 1081 GAAGGGCTGAAATGACCTTCAGGAGGTGAATAA 1119  
 Db 1615 GAAGGGCTGAAATGACCTTCAGGAGGTGAATAA 1653

PUBLISHED	9278803
2 (bases 1 to 13176)	
AUTHORS	Blattner,F.R.
TITLE	Direct Submission
JOURNAL	Submitted (16-JAN-1997) Guy Plunkett III, Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.
REFERENCE	
AUTHORS	3 (bases 1 to 13176)
TITLE	Blattner,F.R.
JOURNAL	Submitted (02-SEP-1997) Guy Plunkett III, Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.
REFERENCE	
AUTHORS	608-263-7459 (bases 1 to 13176)
TITLE	Plunkett,G. III
JOURNAL	Direct Submission
COMMENT	Submitted (13-OCT-1998) Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA. This sequence was determined by the E. coli Genome Project at the University of Wisconsin-Madison (Frederick R. Blattner, director). Supported by NIH grants HG00001 and HG01428 (from the Human Genome Project and NCHGR). The entire sequence was independently determined from E. coli K12 strain MG1655. Predicted open reading frames were determined using GENEmark software, kindly supplied by Mark Borodovsky (Georgia Institute of Technology, Atlanta, GA, le-mail: mark@bambu.gatech.edu). Open reading frames that have been correlated with genetic loci are being annotated with CG Site Nos., unique ID nos. for the genes in the E. coli Genetic Stock Center (CGSC) database at Yale University, kindly supplied by Mary Berlyn. A public version of the database is accessible at <a href="http://www.genetics.wisc.edu">http://www.genetics.wisc.edu</a> . *** The E. coli K12 sequence and its annotations are periodically updated; this is version M54. No ongoing task whose goal is to make the genome sequence more useful by correlating it with other data. Comments to the authors are appreciated. Updated information will be available at the E. coli Genome Project's World Wide Web site ( <a href="http://www.genetics.wisc.edu">http://www.genetics.wisc.edu</a> ). *** The E. coli K12 sequence and sequence changes. Annotation updates: updated gene identifications Riley; added promoters, protein binding sites, and repeated sequences described in reference 1. The unique numeric identifiers beginning with a lowercase "b", assigned to each gene (protein- or RNA-encoding) are how they are designated as gene synonyms instead of labels. This should allow them to be searched for in Entrez as gene names.
FEATURES	
source	/organism="Escherichia coli K12"
promoter	/sub-strain="MG1655"
promoter	/db_xref=taxon:03333"
gene	/note="factor Sigma70; promoter hisS; documented +1 start at 2638662"
promoter	/note="factor Sigma70; complement(86.. .1204)"
CDS	/note="b2515"
promoter	/complement(86.. .1204)
gene	/note="factor Sigma70; complement(86.. .81)"
promoter	/note="factor Sigma70; complement(3786.. .3814)"
CDS	/note="factor Sigma70; predicted +1 start at 2642399"
promoter	/note="b2518"
gene	/note="b2518; complement(3833.. .4264)"
CDS	/note="ndk;"
gene	/EC_number="2.7.4.6"
CDS	/function="enzyme; Purine ribonucleotide biosynthesis"
gene	/note="f143; 100 pct identical to NDK_ECOLI SW: P21233"
promoter	/note="factor Sigma70; hypothetical protein"
CDS	/protein_id="AAC75568_1"
gene	/db_xref="GGID:178863"
promoter	/transl_table=11
gene	/note="f392; 100 pct identical to GCPE_ECOLI SW: P27433"
CDS	/codon_start=1
gene	/transl_table=11
gene	/note="b2515"
promoter	/complement(86.. .1204)
gene	/note="factor Sigma70; complement(86.. .81)"
promoter	/note="b2515"
gene	/note="factor Sigma70; complement(3786.. .3814)"
CDS	/note="b2518"
gene	/note="b2518; complement(3833.. .4264)"
CDS	/note="ndk;"
gene	/EC_number="2.7.4.6"
CDS	/function="enzyme; Purine ribonucleotide biosynthesis"
gene	/note="f143; 100 pct identical to NDK_ECOLI SW: P21233"
promoter	/note="factor Sigma70; hypothetical protein"
CDS	/protein_id="AAC75568_1"
gene	/db_xref="GGID:178866"
promoter	/transl_table=11
gene	/note="f143; 100 pct identical to NDK_ECOLI SW: P21233"
CDS	/note="ndk;"
gene	/note="factor Sigma70; complement(4413.. .6725)"
CDS	/note="b2519"
gene	/note="b2519; complement(4413.. .6725)"
CDS	/note="ppC"
gene	/note="ppC; complement(4413.. .6725)"

```

function="Putative enzyme; Not classified"
note="r770; This 770 aa ORF is 31 pct identical (43 gaps) to 480 residues of an approx.. 720 aa protein PBPF_BACSU
SN: P88050"
/codon_start=1
/codon_table=11
/product="putative peptidoglycan enzyme"
/protein_id="AACI5572_1"
/db_xref="GI: 1788667"
/translation="MPLILTKKGWCWTLAARFLLELAAGWADKLWPLPLHVNPARV
WVQDGIPPLWRKADAWYRPPVTTEDSPRTIEALINYEDFWKWPVNLKINPISVARV
LNRAPEFGTLOGTGAASWALGKSPANLSSPAAMLAVLPQAPSRLPDRPRERAEA
RNKVKLERNVAVQGWSRQPKVPSSEPLTMWLPROMPQLAPLFSRMLGKSDKDTIL
DAGLORRLEELAQNPKVQPLKPSLQVPLDPEGLPDKGKQPLQWVQALQWVQPLVQ
NSIRPSVSKPVYQVGLADELTHPLSPLQVPLDPEGLPDKGKQPLQWVQALQWVQPLV
LYRSNLPVQVLEAYKQVPLQVLEAYKQVPLQVLEAYKQVPLQVLEAYKQVPLQVLE
ATYAFARHGKACKLRLQDPLPPLERPMSSGAAWITRIMDAQPFDPSLPAQPLPVRAPL
AMKTGTSGYRDWATEVNAARYVIGTWGPROGTPVQGIGASAAPVLPNQONTELLS
RSLANLPDPNPNSVTRGVICWQGQSLPREGDNGCNRRIATWLDSQSPPTLPEQG
PYCHDADPLQLOTGVRDGAITKPLGAAETPLQSSGGGERWWELNGEPLPERGRN
VTLHLDKGDYQLEVMDDYQLEVMDDYQLEVMDDYQLEVMDDYQLEVMDDYQLEVM
complement("f16726 .. 11687")
(gene="b2520"
complement("f6726 .. 11687")
(gene="b2520"
/function="orf; Unknown"
/noce="f1653; f1653"
/codon_start=1
/transl_table=11
/product="orf, hypothetical protein"
/protein_id="AACI5573_1"
/db_xref="GI: 1788668"
/translation="MKKLRVAAQMMLALAGCDNNNDAPTVKKDAPSEVTKAASSEN
ASSAKLSPEROKIAQQSRGKVKLTLDLSEVQDGAATVLTISIPLDDQDSRVH
Query Match 100.0%; Score 1119; DB 1; Length 13176;
Best Local Similarity 100.0%; Pred. No. 1.1e-270;
Matches 1119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Query 1 ATGCATAACCAAGGGCTCAATTCAACGTAGAAATCACACGTTATTACGGTGGGATGTG
1204 ATGCATAACCAAGGGCTCAATTCAACGTAGAAATCACACGTTATTACGGTGGGATGTG 1145
Query 61 CCGATTTGGCGATGGGTCTCCATCGCGTAGCTACGCTTACGAGAC 120
1144 CCGATTTGGCGATGGGTCTCCATCGCGTAGCTACGAGAC 1085
121 GTCGAGACAAAGGTCAATCAATCAGGCCGGACGCCAGGGTCAACTCATCACACAGGGTAC 180
1084 GTCGAGACAAAGGTCAATCAATCAGGCCGGACGCCAGGGTCAACTCATCACACAGGGTAC 1025
Query 181 GTATCGGTACCGACGATGGAGGCCAGAGCGTTCACAATCACACAGGGTAC 240
1024 GTATCGGTACCGACGATGGAGGCCAGAGCGTTCACAATCACACAGGGTAC 965
RESULT 7
AV033515 AY033515 1119 bp DNA linear -BCT-26-dBC-20011
LOCUS AY033515
DEFINITION Escherichia coli 1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate synthase (ispG) gene, complete cds.
ACCESSION AY033515
VERSION AY033515.1 GI:17978527
KEYWORDS
SOURCE
ORGANISM Escherichia coli.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
REFERENCE 1 (bases 1 to 1119)
AUTHORS Hecht,S., Eisenreich,W., Adam,P., Amslinger,S., Kis,K., Bacher,A., Arigoni,D. and Rohdich,F.
TITLE Studies on the nonmevalonate pathway to terpenes: The role of the GCPE (IspG) protein
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (26), 14837-14842 (2001)
PUBLISHED 1175/2001
REFERENCE 2 (bases 1 to 1119)
AUTHORS Rohdich,F., Hecht,S., Adam,P., Bacher,A. and Eisenreich,W.
TITLE Direct Submission
JOURNAL Submitted (27-APR-2001) Institut fuer Organische Chemie und Biochemie, Lehrstuhl III, Technische Universitaet Muenchen, Lichtenberg Strasse 4, Garching D-85747, Germany
FEATURES Location/Qualifiers 1 .. 1119
Source 1 .. 1119

```





/transl_table=11	Db	484	GCGCCGATCCGGTCCAGAGATCAAAGTCGGTTCGATATTTGAAATCGCTTCGTATC	425
product="putative dimethyl sulfoxide reductase subunit A"	/protein_id="PAG57632_1"	Qy	CCTTCGGAGGATCACTTCATCGCTGCCGACCTGTCGGTCAGGAATTGAGTT	840
/db_xref="GI-12516916"		Db	424 CGTTCGCGAGGATCACTTCATCGCTGCCGACCTGTCGGTCAGGAATTGAGTT	365
/transl_start="MDTPEKERTTFPOVSRRSFLQATSLITLPISSATAQSPDA				
SPETVAPYADKVPTVPSFSDGCDCIDAHMRDGCVTOUTLPPNEDOPMPIMRACY				
RGRYKRYKVKWPKLRYKVRKRGEEFRISWEATLIDANLRKTIQYGPASR				
YVNGTAVGGTTSGGDAMARLRLKLLGKLELYTHVSVLGNTAAKPTYGVAAQNSM				
DTLUDTKMLWLWHNPNTFTGHNNYYQPKOMKGTRFVWDPRYSDTIVLWLAQDWP				
LUPPTDNALMDAMVYLSENLLKDFTTYGFDENSMPCEGVANPESLVALYFGAK				
DGHKTPWAEKTHYPAQSIRQLADTKEPALIOWGPGORHICGEGTARGESTL				
ASTGNVCKIGKGMAGAAGGSNNRKCVCEDDPMPVQKLTISMWMQADDASVTPQ				
DGKVGDKLSNTRLFLSAGNTLNDPWRQAKLDESKETFTWSDLEMPSA		Qy	ATCGGTACCGTTACCGCCTGGAGAACGCCCTGAGATCATCCTCAATGGACCTT	900
KYDVLKPLPEMSENRWGETGNTASYLNOVLPKQFPERFTDWDLWVAKLGVE		Db	364 ATCGGTACAGTTAACCGCCTGGAGAACGCCCTGAGATATCATCCTCAATGGACCTT	305
AERFSQGRKQKONTEHWTQWQTCLAMPDENPDAFTLQKTRHLRKSAPHAFAINRDP				
QNNPSTPGTKTEIFSKRFDFMDOPDEPILTAFLSHYPAFSGPDEKLTKYPLQIWTWKG				
NRANSTPOTANPMLQEQVOKLWNPQDQDQHNGVSEGSVKLYDORGVSITIPVETPRI				
IPCVVAMQAGAWMWPDAQGDGRGCANCILSTRITALAKGNSHOTMVEKA"				
Db	901	TGATATATCGCTGGCTGATGGCTGAGATGGCCAGGTGGGGCTGTGTTCTACACTCGGCTC	960	
	304	TOGATATATCGCTGGCTGCTGGTGAATGGCCGGGTGGCGCTCGTTCTACACTCGGCTC	245	
Db	961	ACCGGCGCAACAGAAAAGCGGCCCTATGAGAGATGGCGTGGCGCAAGACCGCTCTGGAC	1020	
Qy	244	ACCGGCGCAACAGAAAAGCGCTCTATGAGAGATGGCGTGGCGCAAGACCGCTCTGGAC	185	

	Query	Match	Score	DB	Length
	Local	Similarity		1	
Matches	1103;	Conservative	0;	Mismatches	16;
Qy	1	ATGATAACAGGCTCAATTCAACGTAGAAATCACACGTATTATCAGTGGGAATG	97.7%	DB 1;	Length 11521;
Db	1204	ATGCTAAACAGGCCTCAATTCAACGTAGAAATCACACGTATTATCAGTGGGAATG	98.6%	Pred. No.	3.1e-264;

Db	184	AACACGATATGATCGACCAGCTGGAGCACCGATTCGTGCGAAAGCCAGTCAGCTGGC	125
QY	1081	GAGCGGGTCTGAAATTGACGGTTGAAATAA	1119
Db	124	GAGCGGCTCTGATTGACGTCAGCAGGTGAAATAA	86

DB	1144 CCGATTGGGATGGTGCCTCCATGCCGACAGTCATACCAATACGGGTACAGAGAT	1085
QY	1.21 GTCCGAGCACAGCTAATCAATCAAGGGCTGGAACGGTGGCTGCCCTGATATGTCCT 180	
Db	1084 GTCCGAGCACAGCTAATCAATCAAGGGCTGGAACAGCGTGGCTGAGTCAGCGT 1025	
QY	181 GTATCGTACCGAGCATGGACGCCAGAGCGCTCAAACTCATCAAAAGCAGGTAAAC 240	
Db	1024 GTCTCCGTACCGAGCATGGACGCCAGAGCGCTCAAACTCATCAAAAGCAGGTAAAC 965	
QY	241 GTGCGCGTGTGGCTGACATCCACTCGACTATGCCATTGGCTGAAGACTAGCGGATAAC 300	
Db	964 GTGCGCGTGTGGCTGACATCCACTCGACTATGCCATTGGCTGAAGACTAGCGGATAAC 905	
QY	301 GGCGCTGAGATGCTGGCTTAACCTGCAATTGGAAATGAGAGCGTATTCGAG 360	
Db	904 GGCGCTGATGTCGCTGAAITAACCTGCAATTGGAAATGAGAGCGTATTCGAG 845	
QY	361 GTGGTTGACTGTGCGGCCATAAAACATTCGATCCGGATTGGCTAACCGGAGATC 420	
Db	844 GTGGTTGACTGTGCGGCCATAAAACATTCGATCCGGATTGGCTAACCGGAGATC 785	
QY	421 CTGGAAAGATCTGCAAGAAAGTATGGCAACCGAGGCCGAGGGCTTGCTGGAATCT 480	
Db	784 CTGGAAAGATCTGCAAGAAAGTATGGCAACCGAGGCCGAGGGCTTGCTGGAATCT 725	
QY	481 GCGATGCGCATGTTGATGATCGATGGCAACCGAGGCCGAGGGCTTGCTGGAATCT 540	
Db	724 GCGATGCGCATGTTGATGATCGATGCCCTGACTTGATGATGTTGCTGGAATCT 665	
QY	541 AAAGCGTGTGACGCTCTCTCGTGTGAGTCCTATGTTGCTGCAAAACAGATGGAT 600	
Db	664 AAAGCGTGTGACGCTCTCTCGTGTGAGTCCTATGTTGCTGCAAAACAGATGGAT 605	
QY	601 CAGCGGTGTCATCTGGGNTACCGAAGCCGGTGTGGCCGAGGGGGCAGTAATCTC 660	
Db	604 CAGCGGTGTCATCTGGGNTACCGAAGCCGGTGTGGCCGAGGGGGCAGTAATCTC 545	
QY	661 GCGATGTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720	
Db	544 GCGATGTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 485	
QY	721 GGGGCCATCCSGTCGAGAGATCAAAGTCGGTTGATATTGAAATCGCTGCTGCTG 780	















Kirakawa, M., Makino, K., Masuda, S., Miki, T., Mizobuchi, K., Mori, H., Notomura, K., Nakamura, Y., Nasimoto, H., Nishio, Y., Oshima, T., Saito, N., Sampai, G., Seki, Y., Tagami, H., Takemoto, K., Wada, C., Yamamoto, Y. and Yano, M.

The systematic sequencing of the Escherichia coli genome in Japan  
Unpublished  
REFERENCE  
TITLE  
JOURNAL  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
PROJECT:  
The Japan E.coli genome DNA sequencing project  
Group:  
The Japan E.coli genome DNA sequencing group  
Members: (1995.4 - 1996.3)  
Aiba, H., Baba, T., Fujita, K., Hayashi, K., Honjo, A., Horie, T., Ikemoto, K., Inada, T., Isono, K., Isono, S., Itoh, T., Konai, K., Kasai, H., Kashimoto, K., Kim, S., Kimura, S., Kitagawa, M., Kitakawa, M., Makino, K., Masuda, S., Miki, T., Mizobuchi, K., Mori, H., Motomura, K., Nakamura, Y., Nasimoto, H., Nishio, Y., Oshima, T., Saito, N., Sampai, G., Seki, Y., Tagami, H., Takenoto, K., Wada, C., Yamamoto, Y. and Yano, M.  
Headed by:  
Name: Takashi Horiechi  
Address: National Institute of Basic Biology, Okazaki, 444, Japan  
E-mail: kishor@nibb.ac.jp  
Information operator:  
Name: Hirokada Mori  
Address: NARA Institute of Science and Technology,  
Ibaraki, 630-01, Japan  
E-mail: horie@gtc.aist-nara.ac.jp  
The Japan E. coli genome database  
<http://bsw3.aist-nara.ac.jp>

**FEATURES**

**source**

1. 4775

/organism="Escherichia coli"  
/strain="K12"  
/db\_xref="txon:562"  
/map="56. min"  
/clone="Kohara clone #428"  
/clone\_lib="Kohara lambda miniset library"  
/note="Nucleotide position 2638874-264068 from the initiation site of Tn10 (0 min.). This clone is from Kohara lambda miniset library". complement(1..710)  
/gene="gcpE"  
/complement:<1..710)  
/note="gcpE"  
/note="similar to [PIR Accession Number S23058]"  
/codon\_start=1  
/transl\_table=11  
/product="gcpE protein"  
/protein\_id="BA220919\_1"  
/db\_xref="GI:4433789"  
/translation="MINQAPQRKRTRIYVGNGVPGDADPFTAVOSMTNTTPTDVEATVNQIKALEVRDADIVVSPVTPDAAFKLIKQQVNVPLVADIEHYRTALKVAYGVDCLRINGNIGNERIRVMVCDRDKNIPRIGVNGASLEKDLOQEKGTPPQALLESAKRKHVDLDRMIDQPKVSVASDVFLAVESYRLAKOIQDOPPLHGLTEAGGARSQAVSKAIGIGLLUSEGISPLTR"  
/complement(737..1750)  
/gene="yfgA"  
/complement(737..1750)  
/gene="yfgA"  
/note="similar to [SwissProt Accession Number P27434]"  
/transl\_table=11  
/transl\_table=1

**ORIGIN**

BASE COUNT

a 1092  
a 1304  
a 1228  
a 1151  
t 1304  
t 1228  
t 1151

Query Match

Best Local Similarity

63.4%; Score 710; DB 1; Length 4775;  
Matches 710; Conservative 0; Pred. No. 8.8e+168; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCATAACAGCTCAATTCAACGTTAGAAATCACACCGTATTACGTGGGAATGTC 60  
Db 710 ATGCATAACAGCTCAATTCAACGTTAGAAATCACACCGTATTACGTGGGAATGTC 651

QY 61 CGGATGGGATGGCTCCTCCATCGCGTCACTGAGCAATAGCGTACGACAGAC 120  
Db 650 CGGATGGGATGGCTCCTCCATCGCGTCACTGAGCAATAGCGTACGACAGAC 591

QY 121 GTGGAAGCAAGGGTCAATCAAGC3CTGAAGGCCGTTGGGCTGATATCTCCGT 180  
Db 590 GTGGAAGCAAGGGTCAATCAAGC3CTGAAGGCCGTTGGGCTGATATCTCCGT 531

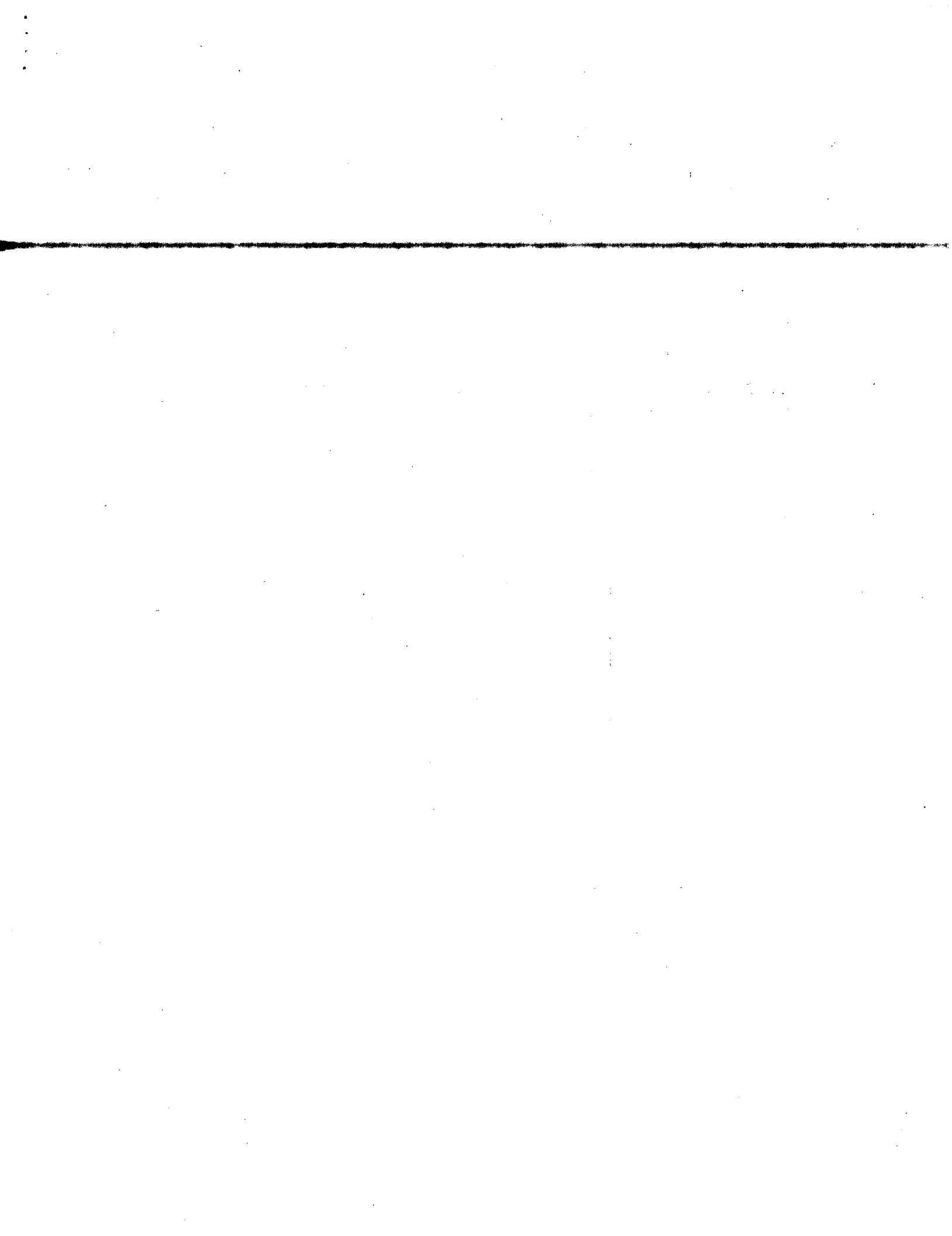
QY	181	GTATCGTACCGACGATGGCCGGAGAACGGTCAACTCATCAAACACAGGTAAAC 240	REFERENCE Mori,H.
Db	530	GTATCGTACCGACGATGGCCGGAGAACGGTCAACTCATCAAACACAGGTAAAC 471	Direct Submission Submitted (23-JAN-1997) Hirotada Mori, NARA Institute of Science and Technology, Res. & Edu. Center for Genetic Info.; 8916-5 Takayama, Ikoma, Nara 630-01, Japan (E-mail: humori@gtc.aist-nara.ac.jp, Tel:81-7437-2-5660, Fax:81-7437-2-5669)
QY	241	GTGCCCTGGTGCATCACATCCACTTCGACTATCGCATTCGCTGAAGAGGGAAATC 300	JOURNAL COMMENT The Japan E. coli genome DNA sequencing project
Db	470	GTGCCCTGGTGCATCACATCCACTTCGACTATCGCATTCGCTGAAGAGGGAAATC 411	Project: Collaboration Information:
QY	301	GCGCTGATGTCGCTGCATCACATCCACTTCGACTATCGCATTCGCTGAAGAGGGAAATC 360	Group: The Japan E. coli genome DNA sequencing group
Db	410	GCGCTGATGTCGCTGCATCACATCCACTTCGACTATCGCATTCGCTGAAGAGGGAAATC 351	Members: (1995.4 - 1996.3) Aiba,H., Baba,T., Fujita,K., Hayashi,K., Bonjo,A., Horuchi,T., Ikemoto,K., Inada,T., Isono,K., Isono,S., Itoh,T., Kawai,H., Kasimoto,K., Kim,S., Kimura,S., Kitagawa,M., Kitakawa,M., Makino,K., Masuda,S., Mikitani,T., Mizobuchi,K., Mori,H., Motomura,K., Nakamura,Y., Nishio,Y., Oshima,T., Saito,N., Sampei,G., Seki,Y., Tegami,H., Takemoto,K., Wada,C., Yamamoto,Y. and Yano,M.
QY	361	GTGGTGACTGGCCGGATAAACATCGATTCGATCGTATGGCGATAACCGGGATG 420	Headed by:
Db	350	GTGGTGACTGGCCGGATAAACATCGATTCGATCGTATGGCGATAACCGGGATG 291	Name: Takashi Horuchi. Address: National Institute of Basic Biology, Okazaki, 444, Japan E-mail: kisnorienib.ac.jp
QY	421	CTGGAAAGATCTGCAGAAAGTGGGACCGACGCCGGGGTTGGTGANCT 480	E-mail: humori@gtc.aist-nara.ac.jp
Db	290	CTGGAAAGATCTGCAGAAAGTGGGACCGACGCCGGGGTTGGTGANCT 231	URL: The Japan E. coli genome database <a href="http://bsw3.aist-nara.ac.jp">http://bsw3.aist-nara.ac.jp</a>
QY	481	GCCATGGTCAGTGTGATCGATCGATCGTCAAGTCAGCTCAGGGT 540	Information operator: Name: Hirotada Mori
Db	230	GCCATGGTCAGTGTGATCGATCGATCGTCAAGTCAGCTCAGGGT 171	Address: NARA Institute of Science and Technology, Ikoma, 630-01, Japan
QY	541	AAAGCCTGTCAGCTCTTCGCTGTGAGTCCTATCGCTTTCGTCGCAACAGATCGT 600	E-mail: humori@gtc.aist-nara.ac.jp
Db	170	AAAGCCTGTCAGCTCTTCGCTGTGAGTCCTATCGCTTTCGTCGCAACAGATCGT 111	FEATURERSOURCE FEATURES source
QY	601	CAGCCGTTGACATGGGGTACCGGAAGCCGGTGGTGGCCGGCGCGCGCAGTAATCC 660	/organism="Escherichia coli" strain="K12" taxon="562"
Db	110	CAGCCGTTGACATGGGGTACCGGAAGCCGGTGGTGGCCGGCAGTAATCC 51	/map="56.5 min" clone="Kohara clone #427" clone_id="Kohara lambda miniset library" note="Kohara lambda miniset library 2619941-2636154 from the initiation site of tRNA (0 min.).-This clone is from Kohara lambda miniset library" <1..1494
QY	661	GCCATGGTTAGTCGCTGCCTGCTGAAAGGATCGCGACAGCTCG 710	/note="similar to [PIR Accession Number A45333]" codon_start=1 codon_end=1 transl_table=11 product="exopolyphosphatase Ppx" protein_id="BA16392_1" db_xref="GI:1805565"
Db	50	GCCATGGTTAGTCGCTGCCTGCTGAAAGGATCGCGACAGCTCG 1	/note="similar to [PIR Accession Number A45333]" codon_start=1 codon_end=1 transl_table=11 protein_id="BA16393_1" db_xref="GI:1805566"
RESULT 15	D90880/C	D90880 16214 bp DNA linear BCT 29-MAY-1997	CDS
LOCUS	D90880	E.coli genomic DNA, Kohara clone #427(56.5-56.9 min.).	
DEFINITION			
ACCESSION	D90880	AB001340	
VERSION	D90880_1	GI:1805564	
KEYWORDS		Complete and shotgun sequencing; AFSK; gcbe; guaa; guab; hisS; xseA;	
SOURCE		Escherichia coli (strain:K12) DNA, clone_lib:Kohara lambda miniset library clone:Kohara clone #427.	
ORGANISM		Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.	
REFERENCE	1	(sites)	
AUTHORS	Yamamoto,Y., Aiba,H., Baba,T., Hayashi,K., Inada,T., Isono,K., Itoh,T., Kimura,S., Kitagawa,M., Makino,K., Mori,H., Mitsuhashi,N., Mizobuchi,H., Mori,S., Nakade,S., Nakamura,Y., Nashimoto,H., Oshima,T., Oyama,S., Saito,N., Sampei,G., Satoh,Y., Sivasundaram,S., Tagami,H., Takahashi,H., Takeda,J., Takemoto,K., Uehara,K., Wada,C., Yamagata,S. and Horiuchi,T.		
TITLE		Construction of a contiguous 874-kb sequence of the Escherichia coli K12 genome corresponding to 50.0-68.8 min on the linkage map and analysis of its sequence features.	
JOURNAL	DNA Res	4 (2), 91-113 (1997)	
MEDLINE	9734980		
REFERENCE	2	(sites)	
AUTHORS	Ikeda,H., Baba,T., Fujita,K., Hayashi,K., Honjo,A., Horiuchi,T., Ikemoto,K., Inada,T., Isono,K., Isono,S., Kanai,K., Kasai,H., Kashimoto,K., Kim,S., Kimura,S., Kitaizawa,M., Kitakawa,M., Makino,K., Masuda,S., Matsuhashi,H., Nishio,Y., Oshima,T., Saito,N., Sampei,G., Seki,Y., Tagami,H., Takemoto,K., Wada,C., Yamamoto,Y. and Yano,M.		
TITLE		The systematic sequencing of the Escherichia coli genome in Japan Unpublished	



Mon May 12 08:44:41 2003

us-09-921-992-3.rge

Search completed: May 10, 2003, 16:14:16  
Job time : 3624 secs



OM nucleic - nucleic search, using sw model  
 Copyright (c) 1993 - 2003 GenCore version 5.1.4\_p5, 4578 CompuGen Ltd.

ried. No. 18 is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
12:				
13: /SIDS2/gcadata/geneseq/geneseqn-emb1/NA1991.DAT:*				AA95478
14: /SIDS2/gcadata/geneseq/geneseqn-emb1/NA1992.DAT:*				ID AA95478 standard; DNA; 1119 BP
15: /SIDS2/gcadata/geneseq/geneseqn-emb1/NA1993.DAT:*				XX
16: /SIDS2/gcadata/geneseq/geneseqn-emb1/NA1994.DAT:*				AC AA95478;
17: /SIDS2/gcadata/geneseq/geneseqn-emb1/NA1995.DAT:*				XX
18: /SIDS2/gcadata/geneseq/geneseqn-emb1/NA1996.DAT:*				DT 27-FEB-2001 (first entry)
19: /SIDS2/gcadata/geneseq/geneseqn-emb1/NA1998.DAT:*				XX
20: /SIDS2/gcadata/geneseq/geneseqn-emb1/NA1999.DAT:*				DE E. coli essential gene gcpe.
21: /SIDS2/gcadata/geneseq/geneseqn-emb1/NA2000.DAT:*				XX
22: /SIDS2/gcadata/geneseq/geneseqn-emb1/NA2001A.DAT:*				KW Bacterial growth; inhibitor; yggf;
23: /SIDS2/gcadata/geneseq/geneseqn-emb1/NA2001B.DAT:*				KW yggf; yieE; yiao; yrdc; yhob; yhob; b1808; yeaA; yagF; b1883; yidD; H. influenza; S. pneumoniae; H. KW M. tuberculosis; antibiotic; ds XX Escherichia coli.
24: /SIDS2/gcadata/geneseq/geneseqn-emb1/NA2002.DAT:*				OS XX PN XX WO20061793-A2.

## ALIGNMENTS



CC in *Saccharomyces cerevisiae*). An antagonist or inhibitor of the expression of an essential gene or of its function provides the key for antibacterial therapy. The invention provides methods for identifying such antagonists or inhibitors. These involve contacting a bacterial cell comprising an essential gene with a candidate antagonist or inhibitor, and testing whether contact leads to cell growth inhibition and/or cell death. The method allows the development of new broad spectrum antibiotics. A conditional mutant of an essential gene can be used to induce a lethal phenotype in bacteria for the analysis of surrogate markers.

SQ

Sequence 1119 BP; 272 A; 279 C; 317 G; 251 T; 0 other;

Query Match 100.0%; Score 1119; DB 21; Length 1119;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCAATAACCGAGGCNCCAATTCACCTGAGAAATCACACGTATTACGTTGGGAATGTG 60  
Db 1 ATGCAATAACCGAGGCNCCAATTCACCTGAGAAATCACACGTATTACGTTGGGAATGTG 60

QY 61 CGGATGGCATGGCTCCATCCGGTCAAGTCATGACCAATACGGTACACAGAC 120  
Db 61 CGGATGGCATGGCTCCATCCGGTCAAGTCATGACCAATACGGTACACAGAC 120

QY 121 GTCGAAAGAACGGTCAATCAAATCAGGGCTGGACGGCTGGCCCTGATATCCCGT 180  
Db 121 GTCGAAAGAACGGTCAATCAAATCAGGGCTGGACGGCTGGCCCTGATATCCCGT 180

QY 181 GTATCGTACCGACGATGGACGGCGAGAAGGGTCAAACTCTATCAACAGCGTTAAC 240  
Db 181 GTATCGTACCGACGATGGACGGCGAGAAGGGTCAAACTCTATCAACAGCGTTAAC 240

QY 241 GTGGTGACTGTGCGCGCATAAACATTCGATCGTATTGGTTAACCGCGGATCG 300  
Db 361 GTGGTGACTGTGCGCGCATAAACATTCGATCGTATTGGTTAACCGCGGATCG 420

QY 421 CTGGAAAAGAGTCTGCAAGAAAGATGGGAGACGAGCGCGCAGCGTGTGAACT 480  
Db 421 CTGGAAAAGAGTCTGCAAGAAAGATGGGAGACGAGCGCGCAGCGTGTGAACT 480

QY 481 GGCATCGTCATGTCATCATCGATCGCTGACTCGATCAAGTCAGCGTG 540  
Db 481 GGCATCGTCATGTCATCATCGATCGCTGACTCGATCAAGTCAGCGTG 540

QY 541 AAAGCGTCTGAGCTTCCTCGCTGTGAGCTTATCGTTGCTGCCAACAGATCGAT 600  
Db 541 AAAGCGTCTGAGCTTCCTCGCTGTGAGCTTATCGTTGCTGCCAACAGATCGAT 600

QY 601 CAGCCGTTGCACTGGGATGACCGAGACGGCGTGGCGCAGGGGAGTAATCC 660  
Db 601 CAGCCGTTGCACTGGGATGACCGAGACGGCGTGGCGCAGGGGAGTAATCC 660

QY 661 GGCATGGTGTAGGTGCGCTGCTGCTGAGGGACATGGGACACGCTGCGCTATGCTG 720  
Db 661 GGCATGGTGTAGGTGCGCTGCTGAGGGACATGGGACACGCTGCGCTATGCTG 720

QY 721 GGGCCGATCCGGTCAACTGAGAGATCAAGTCGGTTGATATTGAAATCCGCTATC 780  
Db 721 GGGCCGATCCGGTCAACTGAGAGATCAAGTCGGTTGATATTGAAATCCGCTATC 780

QY 781 CCTTCGGAGGGATACTTCAGCTGCGGAGGCTGCGTCAAGGAAATTGAGTT 840  
Db 781 CCTTCGGAGGGATACTTCAGCTGCGGAGGCTGCGTCAAGGAAATTGAGTT 840

QY

841 ATCGGTACCGTTAACGGCTGGAGAACCCCTSGAAGATCATCACTCGAVGGACCT 900

Db 841 ATCGGTACCGTTAACGGCTGGAGAACCCCTGGAGATCATCACTCGAVGGACCT 900

QY 901 TCGATTATGGCTGCTGTAATGCCAGGGCTGAGGAGCCCTGGTTCTACCTGGCGTC 960

Db 901 TCGATTATGGCTGCTGTAATGCCAGGGCTGAGGAGCCCTGGTTCTACCTGGCGTC 960

QY 961 ACCGGCGGCAACAGAAAGCGGCCAGGGCTCTAGAAGATGGGTGCCCAGACCGCTGGAC 1020

Db 961 ACCGGCGGCAACAGAAAGCGGCCCTATGAGATGGGTGCCCAGACCGCTGGAC 1020

RESULT 3

ID AAC82653 standard; DNA; 1119 BP.

XX

AC AAC82653;

XX

15-MAR-2001 (first entry)

XX

DE E. coli gpEP DNA.

XX

Isoprenoid; gpEP; yfgB; antimicrobial; transgenic plant; agriculture; KW antiprotoxic; antiparasitic; antiviral; fungicidal; herbicidal; ds.

XX OS Escherichia coli.

XX

W0200072022-A1.

XX

DR 30-NOV-2000.

XX

PR 20-MAY-2000; 2000WO-EP04592.

XX

PR 21-MAY-1999; 99DE-1023567.

XX

PA (JOMA) JOMAA H.

XX

PA Jomaa H.

XX

WPI; 2001-025196/03.

P-PSDB; AAB45692.

PT Incorporating gpEP and yfgB genes into viruses and cells, for increasing isoprenoid content and identifying e.g. antimicrobial agents, comprises using DNA sequences from bacteria or parasites

PT

XX

Claim 3; Page 13-15; 36pp; German.

XX

PS

CC This invention describes a novel method for incorporating gpEP and yfgB genes into viruses and cells for increasing isoprenoid content and identifying e.g. antimicrobial agents, comprises using DNA sequences (I)

CC from the gpEP or yfgB genes of bacteria or parasites or DNA sequences (II) which hybridize to the specified genes or encode a plastid protein with the same biological activity as those encoded by the genes. The invention also describes (1) plant cells containing (I) or (II); (2) transformed plant cells, and transgenic plants regenerated from them,

CC that contain (I) or (II); (3) determining the enzymatic activity of a gpEP protein; or (4) screening compounds (A) that have antiprotoxic, antiparasitic or antiviral activity in humans or animals or antiviral, antiparasitic, fungicidal or herbicidal activity in plants. (I) and (II) are used: (i) to increase the isoprenoid levels in viruses and cells; (ii) for determining the enzymatic activity of gpEP and yfgB proteins; and (iii) to identify compounds that inhibit activity of gpEP, i.e.

CC	potential antibacterial, antimycotic, antiparasitic or antiviral agents for use in humans or animals, or antivirel, antiparasitic, fungicidal or herbicidal agents for agriculture.
CC	Best Local Similarity 100.0%; Score 1119; DB 22; Length 1119;
XX	Matches 1119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
SO	sequence 1119 BP: 272 A; 279 C; 317 G; 251 T; 0 other;
Query Match	100.0%; Score 1119; DB 22; Length 1119;
QY	1 ATGCATAACCGCTCAATCACCTAACGTTAGAAATCACAACCTTATTAGCTGGAAATGTC
QY	1 ATGCATAACCGCTCAATCACCTAACGTTAGAAATCACAACCTTATTAGCTGGAAATGTC
Db	1 ATGCATAACCGCTCAATCACCTAACGTTAGAAATCACAACCTTATTAGCTGGAAATGTC
QY	61 CCGAT <sup>n</sup> GGCATGGCTCCATCGCGTACAGTCATGGACCAATACGGTACGACAGAC 120
Db	61 CCGAT <sup>n</sup> GGCATGGCTCCATCGCGTACAGTCATGGACCAATACGGTACGACAGAC 120
QY	121 GTCGAAGCAAGGTAACTCAATCAAGGCGTGGAAACGCGCTGAACCGTCAACCTAC 60
Db	121 GTCGAAGCAAGGTAACTCAATCAAGGCGTGGAAACGCGCTGAACCGTCAACCTAC 60
QY	121 GTCGAAGCAAGGTAACTCAATCAAGGCGTGGAAACGCGCTGAACCGTCAACCTAC 60
Db	121 GTCGAAGCAAGGTAACTCAATCAAGGCGTGGAAACGCGCTGAACCGTCAACCTAC 60
QY	181 GTATCGTAGCGACGATGGAGCGGGAGAQCCTGAGTTCAGCTTAACACAGGTAAAC 240
Db	181 GTATCGTAGCGACGATGGAGCGGGAGAQCCTGAGTTCAGCTTAACACAGGTAAAC 240
QY	241 GTGGCTGACTGGCTGACATCCACTTGACTATCSCATTGCTGAAGTGCGGAAATC 300
Db	241 GTGGCTGACTGGCTGACATCCACTTGACTATCSCATTGCTGAAGTGCGGAAATC 300
QY	301 GGCGCGATGTCGTTAACCTGGATATCGTAAATGAGAGCGTATTGGAT 360
Db	301 GGCGCGATGTCGTTAACCTGGATATCGTAAATGAGAGCGTATTGGAT 360
QY	421 CTGGAAAAGATCTCGCAAGAAAATGGCATGGCTGAGCTGGTATGGCAATATGAGAGCGTATTGGCT 480
Db	421 CTGGAAAAGATCTCGCAAGAAAATGGCATGGCTGAGCTGGTATGGCAATATGAGAGCGTATTGGCT 480
QY	481 GCCATGCCATGTCATCTCGATCGCTGACTCGCATCGTCAAGTCAGCGTG 540
Db	481 GCCATGCCATGTCATCTCGATCGCTGACTCGCATCGTCAAGTCAGCGTG 540
QY	541 AAAGGTCTGACTCTCCGCGATAAACACATCGATCGCTGAACTTCGATCAAGTCAGCGTG 540
Db	541 AAAGGTCTGACTCTCCGCGATAAACACATCGATCGCTGAACTTCGATCAAGTCAGCGTG 540
QY	601 CAGCGTGTGATCGGGATACCGAAGCGGTGCTGGCGCAGGGGAGTAAATCC 660
Db	601 CAGCGTGTGATCGGGATACCGAAGCGGTGCTGGCGCAGGGGAGTAAATCC 660
QY	661 GCCATTTGTTAGCTCTGGCTGCTGAAAGCATCGGCCACACCGCTGGCGTATCGCT 720
Db	661 GCCATTTGTTAGCTCTGGCTGCTGAAAGCATCGGCCACACCGCTGGCGTATCGCT 720
QY	721 GCGGCCGATCCGGTGAAGAGATCAAGTGGTTTCGATATTGAAATGCTGGTATC 780
Db	721 GCGGCCGATCCGGTGAAGAGATCAAGTGGTTTCGATATTGAAATGCTGGTATC 780
QY	781 CGTTCGGCAGGGATCACTCATCGCCCTGGCTGAGGGCTGGTATC 840
Db	781 CGTTCGGCAGGGATCACTCATCGCCCTGGCTGAGGGCTGGTATC 840
QY	841 ATCGTACGGTTAACGGCTGGAGAACGCCTGGAGAGATATCATCACTCGGATGGACGTT 900
Db	841 ATCGTACGGTTAACGGCTGGAGAACGCCTGGAGAGATATCATCACTCGGATGGACGTT 900
QY	901 TCGATGTTGGCTGGCTGAGGGCTGGCTAACCTGGTACCTGGGCT 960
Db	901 TCGATGTTGGCTGGCTGAGGGCTGGCTAACCTGGTACCTGGGCT 960

CC	used as or primers. The recombinant vectors are used in plant transformation or transfection. gcpE can also act as markers capable of detecting polymorphisms such as single nucleotide polymorphisms (SNPs).
CC	gcpE is also used to determine the level or pattern of expression of the protein. The present sequence is Escherichia coli gcpE gene.
CC	XX
SQ	Sequence 1119 BP; 272 A; 279 C; 317 G; 251 T; 0 other;
Query Match	100.0%; Score 1119; DB 24; Length 1119;
Best Local Similarity	100.0%; Pred. No. 0;
Matches	1119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 ATGCATTAACCGAGCTCAATTCAAGTAGAAATCACACAGTATTAGGTGGGATGTG 60
Db	1 ATGCATTAACCGAGCTCAATTCAAGTAGAAATCACACAGTATTAGGTGGGATGTG 60
Qy	61 CCGATTCGGGTTGGCTCCATCGCGTAGACTCATGACCAAATCGCGTAGACAGAC 120
Db	61 CCGATTCGGGTTGGCTCCATCGCGTAGACTCATGACCAAATCGCGTAGACAGAC 120
Qy	121 GTCGAAGCAACGGTCATCAAAATCAAAAGGCCGTTGAACGCCGCTGATATCGCTCGT 180
Db	121 GTCGAAGCAACGGTCATCAAAATCAAAAGGCCGTTGAACGCCGCTGATATCGCTCGT 180
Qy	181 GATTCGGTACCGACGATGGACCGCGCAGAGGGTCAACATCAACAGCGGTTAAC 240
Db	181 GATTCGGTACCGACGATGGACCGCGCAGAGGGTCAACATCAACAGCGGTTAAC 240
Qy	241 GTGCCGCTGGCGCTGACATCCACTTGACATTCGACTATCGATTCGCTGAAGTAGGGATAAC 300
Db	241 GTGCCGCTGGCGCTGACATCCACTTGACATTCGACTATCGATTCGCTGAAGTAGGGATAAC 300
Qy	301 GCGGTGATGTTGCTGGTATTAAACCTGGCAATATGGTAAAGAGCGTATTCGAT 360
Db	301 GCGGTGATGTTGCTGGTATTAAACCTGGCAATATGGTAAAGAGCGTATTCGAT 360
Qy	361 GGGTGTGACTTGCGGCCGATAAACATTCGATTCGTTAACGCCGATCG 420
Db	361 GGGTGTGACTTGCGGCCGATAAACATTCGATTCGTTAACGCCGATCG 420
Qy	421 CTGGAAAAGACTCTGCAAGAAAAGATATGGCAACGACGCGCAGCGTCTGGATCT 480
Db	421 CTGGAAAAGACTCTGCAAGAAAAGATATGGCAACGACGCGCAGCGTCTGGATCT 480
Qy	481 GCGATGGTCATGTTGATCATCTCGATGCGTGAACCTCGTCAAGTCAGGTG 540
Db	481 GCGATGGTCATGTTGATCATCTCGATGCGTGAACCTCGTCAAGTCAGGTG 540
Qy	541 AAAGCGCTGACTGGCTGCTGAGCTTACCTGGTGGAGCTTACCTGGTGGCAAAAGAGATCGT 600
Db	541 AAAGCGCTGACTGGCTGCTGAGCTTACCTGGTGGAGCTTACCTGGTGGCAAAAGAGATCGT 600
Qy	601 CGGCCGTTGCACTGSGATCCTCGGAGGCCGCTGGCGCAGCGGGCACTAAATCC 660
Db	601 CGGCCGTTGCACTGSGATCCTCGGAGGCCGCTGGCGCAGCGGGCACTAAATCC 660
Qy	661 GCGATTCGGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Db	661 GCGATTCGGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Qy	721 GCGGCCGATCCGCTCGAGAGATCAAGTCGCGTTGATATTGAAATCGTGGGATC 780
Db	721 GCGGCCGATCCGCTCGAGAGATCAAGTCGCGTTGATATTGAAATCGTGGGATC 780
Qy	781 CGTCGGGAGGGATACTCTCGCTGCCGACCGCTGGCGTGGAGATTGATGTT 840
Db	781 CGTCGGGAGGGATACTCTCGCTGCCGACCGCTGGCGTGGAGATTGATGTT 840
Qy	841 ATCGGTACGGTAAAGCGCTGGAGAACGGCTGGAGATAACATCACTCGGTGGACCT 900
Db	841 ATCGGTACGGTAAAGCGCTGGAGAACGGCTGGAGATAACATCACTCGGTGGACCT 900
Qy	901 TCGATATCGCTGGTGAATGCCCAAGCTGGGCTGAGTCAGCTGGCTACCTCGGC 960
Db	901 TCGATATCGCTGGTGAATGCCCAAGCTGGGCTGAGTCAGCTGGCTACCTCGGC 960
Qy	961 ACCGGGGCACACAAAGAACGGCCCTATGAGATGGCTGGGCAAAACGGCTGGAC 1020
Db	961 ACCGGGGCACACAAAGAACGGCCCTATGAGATGGCTGGGCAAAACGGCTGGAC 1020
Qy	1021 AACAAAGATATGATGCCACAGCTGAGCTGGCTGAAGAGCAGCATTCGGCAAGGGCTGGAC 1080
Db	1021 AACAAAGATATGATGCCACAGCTGAGCTGGCTGAAGAGCAGCATTCGGCAAGGGCTGGAC 1080
Qy	1081 GAAGCGCTGCAATGAGCTGAGCTGAGGTTGAAATAAATAA 1119
Db	1081 GAAGCGCTGCAATGAGCTGAGCTGAGGTTGAAATAA 1119
RESULT 5	
ID	AAT42063 standard; DNA; 1830121 BP.
XX	
AC	AAT42063;
XX	
DT	14-SEP-1999 (first entry)
XX	
DE	Haemophilus influenzae complete genome sequence.
XX	
KW	Genome; bacterium; Haemophilus influenzae; computer readable medium; expression modulating fragment; regulation; gene expression; vector;
KW	organism; open reading frame; ORF; ds.
XX	
OS	Haemophilus influenzae.
XX	
PN	W09633276-A1.
XX	
PD	24-OCT-1996.
XX	
PF	22-APR-1996; 96WO-US05320.
XX	
PR	07-JUN-1995; 95US-0487429.
PR	21-APR-1995; 95US-0426787.
PR	07-JUN-1995; 95US-0476102.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC. (UYJO ) UNIV JOHNS HOPKINS.
XX	
PI	Adams MD, Fleischmann RD, Smith HO, Venter JC, White O;
XX	
DR	WI; 1996-485782/48.
XX	
PT	Haemophilus influenzae Rd genome recorded on computer readable medium useful for identifying commercially important nucleic acid fragments by homology searching
PT	
PS	Claim 1; Page 77.2-77.1091; 1291pp; English.
XX	
CC	this sequence represents the complete genome sequence of the bacterium Haemophilus influenzae strain Rd. The invention relates to a computer readable medium (CRM) having recorded upon it the complete H influenzae nucleotide sequence (1), a representative fragment of (1) or a nucleotide sequence at least 99% identical to (1). By providing the full-length genomic sequence in a computer readable form, it is possible to identify commercially important nucleic acid fragments and expression modulating fragments (EMFs) of the Haemophilus genome. The EMFs can be used to regulate the expression of a nucleic acid molecule. Vectors and altered organisms comprising the predicted ORFs can be used to produce any of the polypeptide fragments of the H. influenzae Rd genome.
CC	XX
CC	Sequence 1830121 BP; 567399 A; 350615 C; 347389 G; 564036 T; 682 other;
Query Match	55.6%; Score 622; DB 17; Length 1830121;
Best Local Similarity	74.8%; Pred. No. 3.4e-69;
Matches	794; Conservative 0; Mismatches 265; Indels 3; Caps 1;



QY 755 TCGAATTTGAAATCGCTGATACCGTTCGGAGGATCAACTTCATGCCCTGCAGA 814  
 CC present invention.  
 XX SQ Sequence 640681 BP; 237522 A; 83822 C; 84757 G; 234580 T; 0 other;  
 Query Match 36.1%; Score 403.6.; DB 24; Length 640681;  
 Best Local Similarity 62.3%; Pred. No. 7.2e-06;  
 Matches 653; Conservative 0; Mismatches 394; Indels 3; Gaps 1;  
 Db 2675 CCTGTTGCGTCAGGATTGATGTTACGGTTACCCGCTGGACACGCCGTG 2734  
 QY 815 CCTGTTGCGTCAGGATTGATGTTACGGTTACCCGCTGGACACGCCGTG 874  
 CC 2735 AAGATATCATCTCGATGTTACGGTTACCCGCTGGACACGCCGTG 2794  
 QY 935 AGGCCTGGTCTACTCGATGTTACGGTTACCCGCTGGACACGCCGTG 994  
 Db 2795 AGGCCTGGTCTACTCGATGTTACGGTTACCCGCTGGACACGCCGTG 2854  
 QY 995 ATGGCTGCGCAAAGACCCTGTTGACACACAGATATGATCACCACAGCTGGAGCAGCGA 1054  
 Db 2855 ATGGCTGCGCAAAGACCCTGTTGACACACAGATATGATCACCACAGCTGGAGCAGCGA 2914  
 QY 1055 TTGGCTGCGCAAAGACCCTGTTGACACACAGATATGATCACCACAGCTGGAGCAGCGA 1077  
 Db 2915 TTGGCTGCGCAAAGACCCTGTTGACACACAGATATGATCACCACAGCTGGAGCAGCGA 2937  
 QY 259 ATCCACTTCGACTATCGCATGGCGTGAAGTAGCGGAATACGGCGATGTTCTCGCT 318  
 RESULT 7  
 ABA92787 standard; DNA; 640681 BP.  
 ID ABA92787 standard; DNA; 640681 BP.  
 XX AC ABA92787;  
 XX DT 27-MAR-2002 (first entry)  
 DE Buchnera sp. genomic DNA SEQ ID NO:1.  
 XX KW Buchnera; cockroach-symbiotic bacterium; cockroach extermination;  
 XX KW circular; ds.  
 OS XX DE JP2001292771-A.  
 XX PA (RIKA ) RIKAGAKU KENKYUSHO.  
 DR XX PT WPI; 2002-126043/17.  
 XX PS A genomic DNA of cockroach-symbiotic bacterium -  
 XX Claim 1; Page 16-230; 237pp; Japanese.  
 CC The present invention describes a gene (I) derived from Buchnera sp.  
 CC containing the DNA (a) or (b), (a) has a fully defined base pair.  
 CC sequence selected from a table of sequences found in the Buchnera sp.  
 CC genomic DNA of ABA92787 given in the specification or is a DNA selected  
 CC from complementary DNA sequences, and (b) is a DNA which hybridises with  
 CC the DNA (a) and encodes a protein. Also described are: (1) a recombinant  
 CC vector (II) containing (I); (2) a transformant (III) containing (II);  
 CC (3) a genomic DNA of Buchnera sp. containing the sequence given in  
 CC ABA92787; (4) a plasmid derived from Buchnera sp. containing DNA (c) or  
 CC (d); (5) is a DNA containing a fully defined sequence given in ABA92788  
 CC or ABB92789 and (d) is a plasmid which hybridises with a DNA; and (5) a  
 CC method for the preparation of a protein in which (III) is cultured and  
 CC the expression protein or the objective protein is collected from the  
 CC resultant culture. The DNA is useful for developing agricultural  
 CC chemicals for exterminating cockroaches. The present sequence represents  
 CC the specifically claimed Buchnera sp. genomic DNA sequence, from the  
 CC

Db 31523 AGTCGATTTATGAGACGGAGTRAGACAAAAGAAAAATAAAACGAAGAATTATA 315312  
 QY 1036 GACCAGCTGGAACGACCGATTCTGGCAA 1065  
 Db 315313 GAAGAAATGGAAATTAAATTCGAAAGAAA 315342

RESULT 8  
 AA28536 standard; DNA: 33140 BP.  
 ID AA28536  
 XX  
 AC AA28535;  
 XX  
 DT 04-APR-2001 (first entry)  
 XX  
 DE Genomic fragment #23.

XX  
 KW Genomic library; bacteria; human upper airway; otitis media; sinusitis;  
 bronchopulmonary; endocarditis; meningitis; ss;  
 Moraxella catarrhalis.  
 XX  
 PN WO2000078968-A2.

XX  
 PD 28-DEC-2000.  
 XX  
 PF 16-JUN-2000; 200000-US16649.  
 XX  
 PR 18-JUN-1999; 99US-0140121.  
 XX  
 (INCY-) INCYTE GENOMICS INC.  
 XX  
 PT Lagace RE, Patterson C., Berg KL.  
 XX  
 DR WPI: 2001-041427/05.

XX  
 PS Claim 1; Page 191-199; 545pp; English.

XX  
 CC The present invention relates to a Moraxella catarrhalis genomic library comprising of a combination of 41 nucleic acid molecules (see AAF28534-AAF2854). The library has a number of uses described in the specification e.g. is useful for identifying diagnostic and therapeutic compositions. M. catarrhalis (Branhamella catarrhalis) is a large aerobic, gram-negative diplococcus, normally found among the bacterial flora of human upper airways. M. catarrhalis is known to cause acute, localised infections such as otitis media, sinusitis and bronchopulmonary infection and life-threatening, systemic diseases including endocarditis and meningitis.

XX  
 Sequence 33140 BP; 9472 A; 6496 C; 7384 G; 9787 T; 1 other;

QY 34.2%; Score 382.6; DB 22; Length 33140;  
 Best Local Similarity 59.2%; Pred. No. 2.6e-100;  
 Matches 652; Conservative 0; Mismatches 419; Indels 0; Gaps 0;

QY 1 ATGCATACCGCCGCTCAATTCACTGAGAAATCAACACGTATTACGTTGGAAATGTC 60  
 Db 21735 ATGTCATAACACATCCGATTAAACGCCGCCCCATTACCAAAATCTATGTCGAATGTC 21794

QY 61 CGGATGGCATGGTCTCCATGCCGCTACGCTATGACCATATGGCTACGACAGAC 120  
 Db 21795 GCATCGCCGCGATGACCCATCAGCGTAAAGTATGACAACACTGTGTAT 21854

QY 121 GTCGAACGCAACCGTCATACTCAAACTCAAGGCCGCTGAAACGCCGCTGATATCGTCCGT 180  
 Db 21855 ATTGATGCAACAGTGTGCAACAAATCAGCATGGTGGACGCCAGGTGGACTTGATGCGT 21914

QY 181 GTATCCGTCACCGACGATGCCACGCCGAGAGCGTTCAACTCATCACAAACAGCAGGTAC 240  
 Db 21915 GTATCGACACCAACAGATGGAGAGCGCTCGCTCGCUTTGCGTAAAGAAACGGCTATCC 21974  
 QY 241 GTGCCGCTGGTGGTGCACATCCACTCGACATATCGCATATGCGTAAAGTAGCGGAATAC 300  
 Db 21975 ATCCACATCATGGCATATTGATCACAAAATGCCATITGCAGTGGGGATGTG 22034

QY 301 GGCGTCGATGTGTCMGGTATAACCCCGCAAATCGCTTAATGAGAGCGTATTCGATG 360  
 Db 22035 GGTCGACACTGCCCCTGCGTATCAATCCAGGGTAAATTGGCAATGATCAAAAGTCAAAGAA 22094

QY 361 GTGCTGACTGTGCGCCGTTAAACATCCGATCCGTTATGGCGTACCGCCGATCG 420  
 Db 22095 GTGGTTGCAGGCCACCTCATATGCGCATGGTGTGATCGGTGATGAGTCAGGCTCA 22154

QY 421 CTGGAAAGATCTGCAAGAAAGATGGGAAACCCGACGCCGAGGGTGTGGAATCT 480  
 Db 22155 CTGAAAGGATTTACAAAATAATGGAGAACCGACAGGTGGAGCTG 22214

QY 481 GCCATGCGTCATGTTGATCATCGATGCCGTCACCTTGATGCAACTCACCGTG 540  
 Db 22215 GCGTTACGCCACATTGATTTAGAAACCTTAATTTCAGATAATAAACCTCGATC 22274

QY 541 AAAGCGCTGACGCTCTCCCGTGTGAGCTTATCGTTGCTGGCAAACAGATGAT 600  
 Db 22275 AAAGCCTCCAACTGTTATTTAACCTCTGATGTCATGCGTAACTGACAAATTGAT 22334

QY 601 CAGCGGTGATCTGGGATCACCGAACGGGGTGTGCGAGCGGGCAGTAATCC 660  
 Db 22335 AATCCATGTCATCTGGCTTACCGAACGGGTATATGCTACTGCCAGTCAGATC 22394

QY 661 GCGATTTGGTTAGGTCTCTGCTGCTGCTGAGGCGATCGGCCAACCGCTGCCATGCTG 720  
 Db 22395 GCCATGGCTGGCGGTTGCTACTTGATGCCATGGGATACCATGGTACTCTTG 22454

QY 721 GCGCCGATCCGCTGAGAGTCAGGCTTCATATTGAAATCGCTGGCATTC 780  
 Db 22455 GCGCAGCACCAGAGAGATAAAATTGGCTTGACCATCTAAATCACTGGTAT 22514

QY 781 CGTCGGAGGGATCACTCATGCTGCCGACCGCTCCGTCCGTAGGAATTGATGT 840  
 Db 22515 CGTAGTAATGGCTCAATTAAATGCTTGTCAGGTGTCAGAGTTGATGTG 22574

QY 841 ATGGTAGCTAACGGCTGGAGCAACGCCCTGGAAAGATATCATCACTCCGATGGAGTT 900  
 Db 22575 ATCAAAGTGTGATGAGCTGCTGAGCTGGCTTGGAGGATATTGCTGAGCCACTAGATCTG 22634

QY 901 TCTGATATGGCTGGTGTGATGCGCCAGSTGAGCGCTGCTGGTCTACCTCGCGCTC 960  
 Db 22635 TAGTTATGGCTGTAATGCTATGGCCAGSTGAGCGCCAAAGAGGCCGATCGGTGTC 22694

QY 961 ACCGGCGCAACAGAAAGCGGCCCTATGAGATGCGCTGGCAAAAGACCGTCGAC 1020  
 Db 22695 GTGGGAGCAGGCCAATAGTTGGTTATAACATGTTGCAAAAGCCATCTGATGAT 22754

QY 1021 AACACGATATGATGCCGACGCTGTTGCGGAAGCCGATGCTGCG 1080  
 Db 22755 ACCAAATAATGGTAGATGGAGATTGGCTGGTGTGTCAGTCAGTCAGTC 22814

QY 1081 GAAGCGGTGCGATTGAGCT 1101  
 Db 22815 GATAAACGGCAATGAGATT 22835

RESULT 9  
 AAS9066 standard; cDNA; 1857 BP.  
 XX  
 AC AAS9066;  
 XX  
 DT 13-FEB-2002 (first entry)

XX	DNA encoding novel human diagnostic protein #25870.
KW	Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX	Home sapiens.
XX	WO200175067-A2.
XX	11-OCT-2001.
PD	30-MAR-2001; 2001WO-US08631.
PR	31-MAR-2000; 2000US-0540217.
PR	23-AUG-2000; 2000US-0649167.
XX	(HYSE-) HYSEQ INC.
PA	Drmanac RT, Liu C, Tang YT;
XX	WPI; 2001-639362/73.
DR	P-PSSB; ABG25879.
XX	New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity
PT	Claim 1; SEQ ID NO 25870; 103pp; English.
PS	The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers and for chromosome and gene mapping, and in recombinant production of (III). The poly nucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (III). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (III). (I) and (III) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. As 64197-AAS4564 represent novel human diagnostic coding sequences of the invention.
CC	Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at <a href="http://wipo.int/pub/published_pct_sequences">ftp.wipo.int/pub/published_pct_sequences</a> .
CC	Sequence 1857 BP; 463 A; 472 C; 530 G; 392 T; 0 other;
Query	Best Local Similarity 31.8%; Score 356; DB 23; Length 1857; Matchers 466; Conservative 0; Mismatches 0; Indels 10; Gaps 10;
Matched	644 GCGGGCAGTAAATCCGCCATTGGTTAGTCGTGCTGAGGCAATCGCGACA 703 512 GCGGGCAGTAAATCCGCCATTGGTTAGTCGTGCTGAGGCAATCGCGACA 570 704 CGCTGGGGTATGGCTGGCGGCCATGGCTGAAGAGATAAAGTCGGTTGATATT 763 571 CGCTGGGGTATGGCTGGCGGCCATGGCTGAAGAGATAAAGTCGGTTGATATT 629 764 TGAATTCGCGCTGCGTATGCCCTGCCGACCTGTGC 823 630 TGAATTCGCTG-TATCCGCTGGCGAGGAGTCACATCCCTGCCGACCTGTGC 688 824 GTCAAGAATTGATGTTACGGTAGCTGAGATCAACCTATGCCCTGCCGACCTGTGC 883 689 G-CAGGAATTGATGTTACGGTAGCTGAGATCAACCTATGCCCTGCCGACCTGTGC 746
Result	RESULT 10 ID ABK74401 XX ABK74401 standard; DNA; 1083 BP.
ID	ABK74401;
XX	AC ABK74401;
AC	DE Bacillus licheniformis genomic sequence tag (GST) #1692.
XX	Differential gene expression; genomic sequenced tag; GST; altered culture condition; environmental stress; physiological provocation; ds.
KW	Bacillus licheniformis.
XX	DE Bacillus licheniformis genomic sequence tag (GST) #1692.
KW	Differential gene expression; genomic sequenced tag; GST; altered culture condition; environmental stress; physiological provocation; ds.
XX	OS Bacillus licheniformis.
XX	XX WO200229113-A2.
PN	PD 11-APR-2002.
XX	PF 05-OCT-2001; 2001WO-US31437.
XX	PR 05-OCT-2001; 2000US-0680598.
PR	27-MAR-2001; 2001US-279526P.
XX	PA (NOVO ) NOVOZYMES BIOTECH INC. (NOVO ) NOVOZYMES AS.
XX	PA Berka R, Clausen IG;
PI	XX DR WPI; 2002-416684/44.
XX	PT Monitoring differential expression of several genes in first Bacillus cell relative to expression of same genes in one or more second Bacillus cells, by using substrate containing Bacillus genomic sequenced tag array
PT	PT sequenced tag array
XX	PS Claim 4; SEQ ID NO 1692; 200pp; English.
XX	The invention describes a method of monitoring differential expression of genes in a first Bacillus cell relative to expression of the genes in other Bacillus cells, comprising hybridising labelled nucleic acid probes isolated from Bacillus cells to a substrate containing array of Bacillus genomic sequenced tags (GST), examining the array, and determining relative gene expression by an observed hybridisation reporter signal of a spot in the array. The method is useful for measuring the expression of genes in a first Bacillus cell relative to expression of the same genes in one or more second Bacillus cells. The method is useful for monitoring global expression of several genes from a Bacillus cell, discovering new genes, identifying possible functions of unknown open reading frames and monitoring gene copy number variation and stability. Monitoring changes in expression of genes may be used to provide a representation of the way in which Bacillus cells adapt to changes in culture conditions, environmental stress or other physiological provocation. Extensive follow-up characterisation is unnecessary, when one spot on an array

Page 10

equals one gene or one open reading frame, since sequence information is available. This sequence represents a genomic sequence tag (GST) used in the method of the invention.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at

QY	922	AATGGGCCAAGTGTAGGCCGCTTCTACACTCGGGTCACGGGGCAACAGAAAG	980
Db	921	AACGGTCCCCGAGAACGCCGAATCCGGATCCTGGCACGCCGTAATG	979

Q Sequence 1083 BP; 314 A; 278 C; 265 G; 226 T; 0 other;  
**Query Match** 26.2%; **Score** 293.4; **DB** 24; **Length** 1083;  
**Best Local Similarity** 56.6%; **Pred.** No. 5.3e-75;

ABQ70939  
ID ABQ70939 standard; DNA; 6157 BP.  
XX  
AC  
XX  
ABQ70939;

	Sequence	Conservation	U, 1st nucleic	U, 2nd nucleic	4, 1st nucleic	4, 2nd nucleic	Gaps
22	CAGCTGAAATTCAACAGGT	Y	TATCGTTGGAAATGTGCCATTGCGCATGCTTC	CC	TTT	TTT	81
21	CATCGTACAAACCGCTTGCTTAAGTGGGACCTTAA	b	AGATAGCGGGATAAACGAA	TTA	TTA	TTA	80

DJ  
29-AUG-2002 (first entry)  
XX  
DE  
Listeria monocytogenes 4b contig DNA sequence #881.  
XX  
KW  
Antibacterial; Listeria; food contamination; mutational analysis;

Qy	198	GGACCGGGCAGAGGCTCAACATCAAACAGCAGGTAACTGCAGTGGCCTGGTGCCTGA	257	XX	
Db	3296	GCGTCGAGCAATGCTCTTAGTCGCAATCAAGAAGAGTCATAATCCCTTGTTGCGA	3355	PD	11-APR-2002.
Qy	258	CATCCACTCGACTATCGCATTCGCTGAAGTAGGGAAATCGGGCTGATGTCGCTG	317	XX	
Db	3356	TATCATTGATCCGACTAGTAGCCTTAAGCTTAAGCTTGTGAGGTGACAAATCTCC	3415	PF	04-OCT-2001; 2001WO-FR03061.
Qy	318	TATTAACCTCGGCAATTCGTTAATGCGCTGAAAGTAGGGAAATCGGGCTGATGTCG	317	XX	
Db	3416	GATTATCCTGGTACATCGGTGCGCTGATCGGGTAAAGAGGTATATGCTGCTAA	3475	PR	04-OCT-2000; 2000FR-0012697.
Qy	378	CGATAAACATCCGATCGCTATGGCGTAAACCCGGATCCCTGGAAACAGCTTGTG	437	XX	
Db	3476	AGCAAAATATTCGAATCGCTGATGGGTTATCTGCTGTTAGTAAAGAAATTAT	3535	PA	(INSPIR) INST PASTEUR.
Qy	438	AGAAAGTATGGGACCGGCCGAGGGTTGCTGGAACTGCGATGCTGATGTGA	497	XX	(CNRS ) CNRS CENT NAT RECH SCI.
Db	3536	TCAAATACGGTACCTACTGCTGAAGGAATGGTAAAGTGCACATGATATTA	3595	PI	Kunst F, Glaser P;
Qy	498	TCTATCGATGCCGCACTCGATGATCAACTGCGTAAAGCGCTGAGCTT	557	XX	WPI; 2002-332479/37.
Db	3596	ATTCTCGAAGATTTAGATTTATGATATCATGTTGAAAGCTCTGTGAA	3655	PS	Claim 16; SEQ ID 779; 180pp; French.
Qy	558	CCTCGCTGTGAGCTTATCGTTGCTGGAAACAGATCGATCGTCAGCCGTCATCTGG	617	XX	
Db	3656	TTTACGATGAGCTTATGATTAAGCTAGTCTGCGATTTATCTACANTCTGG	3715	CC	The present invention relates to nucleic acid sequences (AB067188-AB071212) from <i>Listeria</i> sp. The sequences are useful as probes and primers for identification and/or detection of <i>Listeria</i> (e.g. as contaminants in foods, CC
Qy	618	GATCACCGAACCGGGTGGCGAGGGGAGTAAMTCGCCATTGTTAGGCT	677	CC	and for analysis of gene expression. Proteins encoded by the nucleic acid sequences can be used to screen for compounds that modulate gene expression, replication and pathogenicity of <i>Listeria</i> (potential therapeutic agents), also for treating infections by <i>Listeria</i> , and are useful as immunogens in anti- <i>Listeria</i> vaccines.
Db	3716	ATTACGAACTGGGAGGTGCTGGGTTCTGGGGATAAAAGTGGCTGCTT	3775	CC	Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at <a href="ftp://wipo.int/pdb/published_pct_sequences">ftp://wipo.int/pdb/published_pct_sequences</a> .
Qy	678	GCTGCGTCTGAAGGATCAGCGACAGCNGCGCATCCGTGCA	737	CC	
Db	3776	GATACCTAGTTGGCATGGAAATACATTAGGGTACTTGTAGTGTGATCCTGG	3835	CC	
Qy	738	AGAGATCAAGTGGGTTGATATTGAAATCGCGPATCCGTTGGAGGGATCAA	797	CC	
Db	3836	AGAGATAAACGGCCGGGAGGTAAACATCGTTGGCTTCTCGATGCCAT	3895	CC	
Qy	798	CTTCATGCCGCCGACTCTGCTGCTGAGGAAATGANGTATGGTACGGTAAACG	857	CC	
Db	3896	GCTTATCTGCCCTACTGGGTGAATAGAGATGATGTTGATGCTAA	3955	CC	
Qy	858	GCTGGACCAAGCCTGGAGATATCATCCTCCGATGGACCTTCCGATTA	917	CC	
Db	3956	AGTGGAAATTAATCATAGCAACGATTAAAGCCTGATTAAAGTAGCCGTC	4015	CC	
Qy	918	GCTGAAATGGCCGAGGAGGAGCTGCTGTTCTACACTGGCTACCGGGACACGAA	977	CC	
Db	4016	GTCACAAGGTCGGAGAACGCTCGCGAAGCAGATACTGGAAATGCTGG	4075	CC	
Qy	978	AGCGGCTCTATGANGATGG 998		CC	
Db	4076	AGGCCCTCTTGTAGACATGG 4096		CC	
<hr/>					
RESULT 12					
AB067166	AB067966 standard; DNA; 1107 bp.				
ID	AB067966;				
XX	29-AUG-2002 (first entry)				
DE	Listeria monocytogenes EGD DNA sequence #90.				
XX	Antibacterial; Listeria; food contamination; mutational analysis; infection; ds.				
OS	Listeria monocytogenes EGD.				
XX	W0200228891-A2.				

Db	492 AATTCCTCGAGAAGATTTAGATTTATGATPATCATGTTCCTTGAAAGGTTCTGGATGCAA	551	CC	and primers for identification and/or detection of Listeria (e.g. as
Qy	558 CCTCGCTGTGTAATGTTGCTGCAAAACAGAUVCAGCACAGCGTTGATCTGG	617	CC	contaminants in foods, or nutritional analysis) and for analysis of
Db	552 TTAGCAATTGAGCTATGATAAGCTAGTCCGCATTAAATATCCTCTGCATCTGG	611	CC	gene expression. Proteins encoded by the nucleic acid sequences can be
Qy	618 GATCACCGAGCCGGTGGCGCAGGGGGAGTAAATCGCCATGGTAGCT 677	611	CC	used to screen for compounds that modulate gene expression, replication
Db	612 ATATACAGATCTGTAACATTGCTGGAGATAAAAGTGCTGCTGGTTAGGAC	671	CC	and pathogenicity of Listeria (potential therapeutic agents), also for
Qy	678 GCTGCTGCTGAGGCATCGGCCACCCCTGCCGTATGCCGCTGGCGGACCGGTGA	737	CC	treating infections by Listeria, and are useful as immunogens in
Db	672 GATACTCAGTTGGCATTTGAAATACATTACAGATATCTTGTAGTGTGATCTGTGGA	731	CC	anti-Listeria vaccines.
Qy	738 AGAGATCAAGTGGTTGATTTGAAATGCTGGTATCGTGGAGGATCAA	797	CC	Note: The sequence data for this patent did not form part
Db	732 AGAAATAAACAGTGCCGGAAAGTTAAATCATTGCTCTTCCGAAACCGGAAT	791	CC	of the printed specification, but was obtained in electronic format
Qy	798 CTCATCGCTGCGACGCTGAGATTCATCGCTGAGATAGATGATTATCGTATCCCTAATGA	851	CC	directly from WIPO at <a href="http://ftp.wipo.int/pub/published_pct_sequences">ftp.wipo.int/pub/published_pct_sequences</a> .
Db	792 GCTTATCTCTGCTACTGCGACTACCTGCGCTGAGATAGATGATTATCGTATCCCTAATGA	851	XX	Sequence 1107 BP; 348 A; 178 C; 255 G; 326 T; 0 other;
Qy	858 CGCGGAGAACGCTGGAGATACTCGATGCGATGGATTGATGTTACGGTGGATGCTG 911	911	Query Match	22.5%; Score 251.4; DB 24; Length 1107;
Db	852 AGTGGAAATTAAATAGAGATGAGTTGGATCGGATTAAGTGGCTGCTG 911	911	Best Local Similarity	53.5%; Pred. No. 8; 8e-63;
Db	918 GCTGAATGCCGCCGGTGGCTGGTTCTACACTCGCGTACCCGGGACAAGAA	977	Matches	525; Conservative
Db	912 GTCACACGGCCCTGGAGAGCTCGCGAACGGCATATCGGAATGCTGGTCAACGGAGA	971	Mismatches	0; Indels
Qy	978 AACGGCTCTATGAGGG	998	Db	72 GGATTAACCTATCCAACAGCTACTACATACAGACATGAGCTGAGCAGACAG 131
Db	972 AGCCCTCTTTRAGACATGG	992	Qy	18 AATTCACAGTAGAAATCACAGCTATTTACGTTGGGAATGCGGATGGCGATGGTGC 77
RESULT 13			Db	12 AATTTTCGCAAACACTCGCCAGTCAGTCAGTACCCANTACGGTACACAGAGCTTAACTTGTGTTAATGCTG 71
ID	ABO6993		Db	78 TCCCATCGCGTACAGTCAGTACCCANTACGGTACACAGAGCTTAACTTGTGTTAATGCTG 77
XX	ABO6993 standard; DNA; 30107 BP.		Db	132 AGAAATTCACCGATTAGAGAAGCTGGTTGTCAGATGTGGAGTTCTGTCTGTGATGA 191
AC			Qy	198 GGACGGCCAGAACGGTCAACTCATCAACAGCAGGTTAACGTGCGCTGGGTGA 257
XX			Db	192 ACGTGCAAGGAATGTCCTAGTGGCATCAAGAAAAGATCATATTCGCTGTGCGA 251
DT	29-AUG-2002 (first entry)		Qy	258 CATCCACTCTGACTATGCGATGGCGTAAAGTAGCGGATACGGGATCGGCCTGAGTGTGAGCTG 317
XX			Db	252 TATCCATTGATTAACCGACTAACGACTAACGACTAAAGCTATGAGCTGAGGTTGACAAATCG 311
DE	Listeria monocytogenes EGDe DNA sequence #205.		Qy	318 TATTAACCTGGCAATATCGGATATGAGAGCGTATTGCGATGGTGTGACTGCGCG 377
KW	Antibacterial; Listeria; food contamination; mutational analysis;		Db	312 GATTAATCCTGGTAAACATGGTGCCTGGATCGGGTTGAAAGATGTGCTAA 371
KW	infection; Qs.		Qy	378 CGTAAACACATCCGATCGCTGTTAGGGCTAACCGGGATGCTGGAAAGATCTGCA 437
XX			Db	372 AGCAAAATAATTGCAATCGTATTGGGTTATGCTGGTAGTTAGAAAAGAAATTAT 431
OS	Listeria monocytogenes EGDe.		Qy	438 AGAAAGTGGAAACGCCGCCGCGCGTGTGCTGAGATCTGCCATGCCATG 497
XX	W0200228891-A2.		Db	432 TCAAAATATGGTACCCACTCTACTGCTGAGGAATGTTGAGAATGCACTTGC 491
PN	04-OCT-2000; 2000FR-0012697.		Qy	498 TCATCTCGATCGCTGAACTTCGATCATGTCAGTCAGTCAGCGGAAAGCGCTGACTCT 557
XX	11-APR-2002.		Db	492 AATTCTCGAGATTAGATTTATGATATCATGTTCTTGAAGGCTCTGATGTGAA 551
XX	04-OCT-2001; 2001WO-FR03061.		Qy	558 CCTCGCTGTGACTATGCTGTTGGCAACACATCGATCACCCGTCACCTGG 617
PR	New genomic sequences from Listeria species, useful for detection, treatment and prevention of infection, also related polypeptides, antibodies and modulators	-	Db	552 TTAGCAATTGAGCTATGATAAGCTGCGATTTAATCTCTCTGCAATCTGG 611
PA	(INSP ) INST PASTEUR.		Qy	618 GATCACCGAGCCGGTGGCGCCAGGGGGAGTAAATCGCCATGGTTAGGCT 677
PA	(CNRS ) CNRS CENT NAT RECH SCI.		Db	612 ATATACAGAATCTGTCACAAATTGCTGGAGAATAAAGTGTGCTGCTGTTAGGAC 671
XX	Kunst F, Glaser P;		Qy	678 GCTGCTGCTGAGCTGCGACAGCTCCGCTGGCGATTCCTGCGCCATCCGTCGA 737
DR	WPI; 2002-332479/37.		Db	672 GATACTCAGTTGGCATGGAAATACATTAGGATATCTTGTAGTGTGCTGATCTGG 731
XX	The present invention relates to nucleic acid sequences (AB067188-WO20021212) from Listeria sp. The sequences are useful as probes		Qy	738 AGAGATCAAACGGGTTGATTTGAAATGCGCGCATCCGTTGGGGGATCAA 797
PT			Db	732 AGAAATAAACAGTGCCGGGAGTTAAACATTTGCTCTCTGCAATCCGCAAT 791
XX	Claim 16; SEQ ID 2805; 180pp; French.		Qy	798 CTTCATGCCCTCCGACCCGTTGCTGCAAGTATGCTGTTAGGTTAACGC 857
CC			Db	792 GCTTATCTCGCCACTTCGGTGTGAAATAGAGATGTTAATGCTGATATGCTAATGA 851



Db 25418 ggcAAAGATCTGCAAGCATGGCGAGGTCA 25391  
 Qy 229 CAGCAGTTAACGTGCCCTGGCTGACATCCACTCGCATTCGCCTGAAA 288  
 Db 256 CGCACGCGACATCCGGTAACTCGCGGACATACATTCAGCGCGCTACATATCGCC 315  
**RESULT 15**  
 ID AAH52056  
 AAH52056 standard; DNA; 1164 BP.  
 AC XX  
 XX  
 DT 04-SEP-2001 (first entry)  
 DE Mycobacterium tuberculosis potential drug target gene SEQ ID 110.  
 XX  
 KW Drug target; growth; organism viability; characterisation; ds.  
 XX  
 OS Mycobacterium tuberculosis.  
 PN WO2001353117-A1.  
 XX  
 PD 17-MAY-2001.  
 XX  
 PR 13-NOV-2000; 2000WO-US31152.  
 XX  
 PR 12-NOV-1999; 99US-0165086.  
 PR 12-NOV-1999; 99US-0165124.  
 PR 01-FEB-2000; 2000US-0179531.  
 PA (REGC ) UNIV CALIFORNIA.  
 XX  
 PT Eisenberg D, Rotstein SH, Marcotte EM;  
 XX  
 DR WPI; 2001-329193/24.  
 PT-PSDB; AAG81205.  
 XX  
 Identifying nucleotide or polypeptide sequence for use as drug target, involves providing algorithm that analyzes a functional relationship between nucleotide or polypeptide sequences, and comparing the sequences disclosure; Page 124-125; 207pp; English.  
 XX  
 This invention relates to a method for identifying a nucleotide or polypeptide sequence that may be a drug target, or essential for growth or viability of an organism. Polynucleotide sequences AAI51947 - AAI52092 represent DNA encoding proteins AAG81096 - AAG81241, Mycobacterium tuberculosis proteins which are potential drug targets. The DNA and protein sequences are used to illustrate the method of the invention. The method involves providing an unknown nucleotide or polypeptide sequences, and comparing it to a number of sequences along with at least one algorithm capable of analysing a functional relationship between nucleotide and polypeptide sequences. The method is useful for characterising the function of nucleic acids and polypeptides that may be useful as a target for a drug or essential for the growth or viability of an organism.  
 XX  
 Sequence 1164 BP; 203 A; 357 C; 402 G; 202 T; 0 other;  
 SQ Query Match 21.8%; Score 243.8; DB 22; Length 1164;  
 Best Local Similarity 53.5%; Pred. No. 1.5e-60; Mismatches 534; Conservative 0; Mismatches 462; Indels 3; Gaps 1; Matches 136; Search completed: May 10, 2003, 15:48:52  
 Job time : 2434 secs  
 Db 316 GGCATCGACGCCGATGCGATGTCGGCGGGCGGTCACCGATTCAGCGCTTGGCGCT 108  
 Qy 289 GTAGCGGAATAAGGGCGTGTGATGTCCTCGGCTAAGCGGATTAACCTSGCAATATCGCGCTACATATCGCC 315  
 Db 346 GAGCGTATTGCGATGGTGTGACTGTGCGCGGATAAAACATTCGATTCGATTCGTAATWGGC 405  
 Db 376 GCGCGGTGGGAGGGCGCAAGGGCGGGTGCGGCCGATCCGATCCGATCCGATTCGGA 435  
 Db 406 GTAAACCCGGATCGCTGGAAAAGATCTGCAAGAAAGATGGCGAACCGACGCGCAG 465  
 QY 436 GTCACACGGCTCGCTGGACAAACGCTTCATGGAGAGATGGCAAGGCGAC 495  
 Db 466 GGGTGTGGAATCTGCCATGGCTATGTGATCATCTCGATCGCTGACTTCGATCAG 525  
 Db 496 GCGCTGTGTGACTCGCGCTGCGGGCTGGAGAGCTTGCTGTTTCGAGGASCATGGCTTGGTAC 555  
 Db 526 TCCAAGTCAGGTGAAGCGCTGCTGACGCTCTTCGCTGCTGTTGAGCTTATGTTGCTG 585  
 Db 556 ATCAAGATCAGGTCAGCACACGACGACGGGGGGTGATGTCGCGCCCTACAGGAGT 615  
 Db 586 GCAAACAGATCGATCAGCGCTGCACTCTGGATGACCGAGCCGGTGGCTGGCCAC 645  
 QY 616 GTGTCACCGTGGCACTACCACTGACCTCGCTGGTCACTGGAGCCCTGCTTCCAG 675  
 Db 646 GGGCAGTAAATGCCATTGGTTAGGTCTGGCTGCTGCTGCTGCTGAGAGATCAAGTCGGTTGATTTG 705  
 Db 676 GGCACCACTAACGCCATCGCGGTTGGCTTGGCGCTGCTGCGGGCATAGGCAC 735  
 QY 706 CTGGCGTATCGTGGGGCATCCGGTCTGGCTGAGAGATCAAGTCGGTTGATTTG 765  
 PT 736 ATCGGGGTGCGTGTGCTGGCCGGCTGGAGAAGTCAGTGGGCAATGGGACAC 795  
 Db 796 GAGTCGTGAACTCTGCCCGCGCTGCTGAGATCTGTCGCTGCGCTCG 855  
 Db 826 CAGGAATTGATGTTATCGTACGGTAAACGGCTGAGAACGCAACGCGCTGGAGATATC 885  
 Db 856 GGCACAGTCGACCTACACCTGGCCACAGAGGTAACCGCCGGCTGGATGTCCTGAT 915  
 QY 886 ACTCGATGAGCTTTCGATATCGCTGCGGGTGAATGCCAGGGAGGCGCTGGT 945  
 Db 916 GAGCGGTTGGGGCTGGCTGAGGGGGTGTGTCATGGTGGGTTGAGCACGTGAG 975  
 QY 946 TCTACACTCGGGCTCACGGGGCACAGAAAAGGCCCTATGAGATGGCTGCG 1005  
 Db 976 GCGACCTGGCTGGCTGGCTGGCAACGGCAAGGTCAGATTTGTTGAGGAGT 1035  
 QY 1006 AAAGACCGTCTGACAACAGATATGATGACCGAGCTG 1044  
 Db 1036 ATCAGACCGTCTGACAACAGATCTGGAGAGCTG 1074

GenCore version 5.1.4\_p5\_4578  
copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model1

Run on: May 10, 2003, 15:05:07 ; Search time 1889 Seconds  
(without alignments)  
9593.833 Million cell updates/sec

Searched: US-09-921-992-3  
Perfect score: 1119  
Sequence: 1 atgataaacaggctccaaat.....ttcagcaggttggaaaataa 1119

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*

1: em\_estiba:\*

2: em\_lesthum:\*

3: em\_lestin:\*

4: em\_lestiu:\*

5: em\_lestov:\*

6: em\_lestpl:\*

7: em\_lestro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_hcc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: gb\_gss:\*

18: em\_gss\_hun:\*

19: em\_gss\_inv:\*

20: em\_gss\_pth:\*

21: em\_gss\_vrt:\*

22: em\_gss\_fun:\*

23: em\_gss\_main:\*

24: em\_gss\_mus:\*

25: em\_gss\_other:\*

26: em\_gss\_pro:\*

27: em\_gss\_rnd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

**SUMMARIES**

Result No.	Score	Query Match Length	DB ID	Description
C 1	81.8	7.3	351 14 BU004160	BU004160 OGG37N12.
C 2	81.6	7.3	720 17 BH235010	BH235010 PH_05.X_S
C 3	67.8	6.1	448 10 AV644003	AV644003 AV644003
C 4	67.4	6.0	482 10 AV642998	AV642998 AV642998
C 5	67.4	6.0	502 10 AV626792	AV626792 AV626792
C 6	67.4	6.0	512 10 AV644476	AV644476 AV644476

**ALIGNMENTS**

RESULT 1	LOCUS	DEFINITION	VERSION	KEYWORDS	SOURCE	ORGANISM
BU004160/c	BU004160	OGG37N12.yg.ab1	351 bp	mRNA linear	EST 22-AUG-2002	Lactuca sativa
						CDNA clone
						OGG37N12, mRNA sequence.
						BU004160
						BU004160.1 GI:22438555
						EST.
						Lactuca sativa
						Lactuca sativa
						Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids II; Asterales; Asteraceae; Lactuceae; Lactucae.
						(bases 1 to 351)
						1 (bases 1 to 351)
						Kozik,A., Michelmore, R.W., Knapp,S., Matvienko,M., Rieseberg,L., Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J., Ellison,P., Kolkman,J., Stabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z., Church,S., Jackson,I. and Bradford,K.
						Lettuce and Sunflower BSRs from the Compositae Genome Project
						<a href="http://compgenomics.ucdavis.edu/">http://compgenomics.ucdavis.edu/</a>
						Unpublished (2002)
						Contact: Alexander Kozik [R.W.Michelmore]
						Department of Vegetable Crops, R.W.Michelmore Lab
						University of California at Davis (UCD)
						Ashmond Hall, UCD, Davis, CA 95616, USA
						Tel: 1-(530)-742-1742
						Fax: 1-(530)-752-9659
						Email: <a href="mailto:akozik@atgc.org">akozik@atgc.org</a> [ <a href="mailto:michelmore@vegemail.ucdavis.edu">michelmore@vegmail.ucdavis.edu</a> ]

FEATURES	singleton; see <a href="http://cgpdb.ucdavis.edu/">http://cgpdb.ucdavis.edu/</a> for details.
source	Location/Qualifiers
1. .351	1. .351
/organism="Lactuca sativa"	
/clone="QGG7N12"	
/lab_vector="E. coli"	
/clone_1.lib="Q_EFGHJ lettuce serriola"	
/db_xref="txon:4236"	
/clone_1.lib="Q_EFGHJ lettuce serriola"	
/note="vector: pRCDNASFiAB; The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made into four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at <a href="http://cgpdb.ucdavis.edu/">http://cgpdb.ucdavis.edu/</a>	
BASE COUNT	62 a 114 c 115 g 59 t 1 others
ORIGIN	
Query Match 7.3%; Score 81.8; DB 14; Length 351; Best Local Similarity 53.1%; Pred. No. 4.8e-14; Mismatches 173; Conservative 0; Indels 0; Gaps 0;	
QY 429 AGATCTGAAAGAAAGTATGGCAACGCCAGCGCAGCGGTCTGGAAATCTGCCATCGC 488	
Db 332 AGCTCTGCCAACAGTGGCCGCCACGCCGCTGAGCGGTCTGGCTCCCTCTCAA 273	
QY 489 TCGTGTGATCATCTCGATCGCCGAACTTCGATCAAGTCAGCGTSAAGCGTC 548	
Db 272 GGAGCGGCCATGTCGAGATCTGGCTTCCCGCCCTCAAGATCTGGTCAAGCACCA 213	
QY 549 TGGCTCTCCRGCTGTGAGCTTAACGTTGCTGCAAAACAGATCGATCAGCGTT 608	
Db 212 CGCCGGGACCATGGTGGAGACTACCGACTCTCGCTCCAAAGGAGCTGTCCCT 153	
QY 609 GCATCTGGGACACGGAGCCGAGCGGGTGGCCGAGCGGGAGTAMATCCGCATGG 668	
Db 152 GCACCTGGAGACTGAGGAGCTGGCCCTGCAAGCCCCATCACTCTGGTGGC 93	
QY 669 TTTAGGCTGCTCTGTGTAAGGCATGGCGGACCGCTGGCTATCGCTGGGGCGA 728	
Db 92 ATTCGGTGCCTCTCGCCGAAGGCATGGCTTGCTCTCNGGCC 33	
QY 729 TCGGTGAAGAGATCAAAAGTGGTT 754	
Db 32 TCTGTGGAGGAGGTGAAGTGGGT 7	
RESULT 2	
BH235010 LOCUS BH235010 720 bp DNA linear GSS 01-JAN-2002	
DEFINITION BH_05_x Spiroplasma kunkelii H Spiroplasma kunkelii genomic clone	
ACCESSION PH_05_x, DNA sequence.	
VERSION BH235010.1 GI:18030478	
KEYWORDS GSS.	
SOURCE Spiroplasma kunkelii	
ORGANISM Spiroplasma kunkelii	
Bacteria; Firmicutes; Mollicutes; Entomoplasmatales; Spiroplasmataceae; Spiroplasma.	
REFERENCE 1 (bases 1 to 720)	
AUTHORS Hogenhout,S.A.	
JOURNAL Unpublished (2001)	
COMMENT Contact: Erika Asamizu	
THE FIRST LABORATORY FOR PLANT GENE RESEARCH	
KAZUSA DNA RESEARCH INSTITUTE	
YANA 1532-3, KISARazu, Chiba 292-0812, Japan	
E-mail: asamizukazusa.or.jp, URL: <a href="http://www.kazusa.or.jp/en/plant/">http://www.kazusa.or.jp/en/plant/</a>	
FEATURES Location/Qualifiers	
source	Location/Qualifiers
1. .448	1. .448
/organism="Chlamydomonas reinhardtii"	
/strain="C9"	
/db_xref="taxon:3055"	
/clone="RHCT080502_r"	
/clone_1.lib="Chlamydomonas reinhardtii 5% CO2"	
FEATURES Location/Qualifiers	
source	Location/Qualifiers
1. .720	1. .720
/organism="Spiroplasma kunkelii"	
/strain="M2"	
/db_xref="txon:47834"	
/clone="PH_05.x"	
BASE COUNT	255 a 96 c 132 g 233 t 4 others
ORIGIN	
Query Match 7.3%; Score 81.6; DB 17; Length 720; Best Local Similarity 53.4%; Pred. No. 8.5e-14; Mismatches 191; Conservative 0; Indels 1; Gaps 1; Gaps 1;	
QY 572 CTATCGTTGCTGGCAAAACAGATCGATCAGCGTTCAGCGTTCATGGGATC,CCGAGGCC 631	
Db 35 CTATACCTAGTAGTAAAGAATGAAATTACCCCTCATCTAGGGAT-GCTGAAGCT 93	
QY 632 GTGGTGCGGCAGCGGGCAAGTAAATGCCATGGTTAGCTGCTGCTGCTG 691	
Db 94 GTACTCATCATACGGAAACAUAAAATGAGTAGTGGTTGACCGCTCTTTAATG 153	
QY 692 GCATGGGCCACAGCTGGCGTATCGCTGGCGCCATCGCGTCAAGGATCAAGTC 751	
Db 154 GTATGGTGTATGATCCGAATTAGTTATCAACTGATCATACTAGCAGAGGGAGTT 213	
QY 752 GTTCGATATTGAAATGCTGGTATCCGGTTCGCGAGGGATCAACTCATGCC 811	
Db 214 GTAAACGGAATGTTAAATCTATAGACTTTATGATACTATGTTGATATTGCTG 273	
QY 812 CGACCTGTTGCGTCAGGATTGATGTTACCGTTAACGCCTGGACAAAGGCC 871	
Db 274 CAACATGTTGCTGTTAGATATGACCTTTCGGTGTGAARAAGAATTAGAAATA 333	
QY 872 TGGAGATATCATCTCGCATGGCTGGACGTTCTGCTTATTCGGCTGGGGTAATGGGCC 929	
Db 334 CAAAGATTAATCCTTAAATGCCATTAGTGTGTGTGTTAATGGTCC 391	
RESULT 3	
AV644003 LOCUS AV644003 448 bp mRNA linear EST 15-DEC-2000	
DEFINITION AV644003 chlamydomonas reinhardtii 5% CO2 chlamydomonas reinhardtii mRNA sequence.	
ACCESSION AV644003	
VERSION AV644003.1 GI:10787331	
KEYWORDS EST.	
SOURCE Chlamydomonas reinhardtii	
ORGANISM Chlamydomonas reinhardtii	
Bukaryote; Viridiplantae; Chlorophytta; chlorophyceae; volvocales; chlamydomonadaceae; chlamydomonas.	
REFERENCE 1 (bases 1 to 448)	
AUTHORS Asamizu,E., Miura,K., Kuchio,K., Inoue,Y., Fukuzawa,H., Ohya,M., Nakamura,Y. and Tabata,S.	
TITLE Generation of expressed sequence tags from low-CO2 and high-CO2 adapted cells of Chlamydomonas reinhardtii	
JOURNAL DNA Res. 7 (5), 305-307 (2000)	
MEDLINE 2053064	
COMMENT The First Laboratory for Plant Gene Research	
KAZUSA DNA Research Institute	
YANA 1532-3, Kisarazu, Chiba 292-0812, Japan	
E-mail: asamizukazusa.or.jp, URL: <a href="http://www.kazusa.or.jp/en/plant/">http://www.kazusa.or.jp/en/plant/</a>	
FEATURES Location/Qualifiers	
source	Location/Qualifiers
1. .448	1. .448
/organism="Chlamydomonas reinhardtii"	
/strain="C9"	
/db_xref="taxon:3055"	
/clone="RHCT080502_r"	
/clone_1.lib="Chlamydomonas reinhardtii 5% CO2"	

	/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XbaI; The cDNA library was constructed from cells cultured in a medium with bubbling air containing 5% carbon dioxide"
BASE COUNT	106 a 142 c 132 g 68 t
ORIGIN	
Query Match	6.1%; Score 67.8; DB 10; Length 448;
Best Local Similarity	56.6%; Pred. No. 1.2e-09;
Matches	175; Conservative 0; Mismatches 116; Indels 18; Gaps 2;
Db	97 TCGGAAGGTGCCGATGGCGATGTCCTCCATGCCCTACAGTCAGGAAATACCC 109
Qy	50 TTGGAAATGCGCATGGCGATGTCCTCCATGCCCTACAGTCAGGAAATACCC 109
Db	97 TCGGAAGGTGCCGATGGCGATGTCCTCCATGCCCTACAGTCAGGAAATACCC 156
Qy	110 GTACGACAGCAGTCGCAAGCACGGTCAATCAATCAGGGCTGACCGTTGGCG 169
Db	157 ACACCGCAATGTCAGCTGACCGTGGACCGACCG 216
Qy	170 ATATGGTCCGTTATCCGACGATGGACGGCAGAGCGTCA----- 217
Db	217 ACTATGGTCCGATCACAGTCGAGGCAGAGGAGGCTGAGGGCGTCAAGAATTCGG 276
Qy	218 --AACTCATCACAGCAGTTAACGTGCGCTGGTGTGACATCCACTTCGACTATC 274
Db	277 AGCAGCTGTCAGGACAAGTACGAGTGGCCCTGGTGGCGAGATCCACTCCAGCCC 336
Qy	275 GCATGGCGTGAAGTAGCGGAATAGCGGAATAGGGCTCGATGTTGGGTATTAAACCTGCAATA 334
Db	337 CGGTGGCAGATGGGGCG--ACGCCCTTGAGAAGATGCCGTTGAGACCCGCGCAACT 393
Qy	335 TCGGATGAA 345
Db	394 TCGCTGACGGA 404
RESULT 4	
AV642898	AV642898 482 bp mRNA linear EST 15-DEC-2000
LCCUS	Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii
DEFINITION	AV642898 Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii
ACCESSION	AV642898
VERSION	AV642898.1 GI:10786226
KEYWORDS	EST.
SOURCE	Chlamydomonas reinhardtii.
ORGANISM	Chlamydomonas reinhardtii.
REFERENCE	1 (bases 1 to 482)
AUTHORS	Asanizu,E., Murakami,K., Kuchio,K., Inoue,Y., Fukuzawa,H., Ohyama,K., Nakamura,Y. and Tabata,S.
TITLE	Generation of expressed sequence tags from low-CO2 and high-CO2 adapted cells of Chlamydomonas reinhardtii
JOURNAL	DNA Res. 7 (5), 305-307 (2000)
MEDLINE	20539644
COMMENT	Contact: Erika Asanizu
The First Laboratory for Plant Gene Research	
Kazusa DNA Research Institute	
Yana 1533-3, Kisarazu, Chiba 292-0812, Japan	
Email: asanizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.	
FEATURES	location/qualifiers
source	1. .502
/organism="Chlamydomonas reinhardtii"	
/strain="C9"	
/db_xref="ttaxon:3055"	
/clone="JCL015033_r"	
/clone_110=Chlamydomonas reinhardtii 5% to 0.04% CO2"	
/note="Vector: pBluescriptSK-; Site_1: EcoRI; Site_2: XbaI; The cDNA library was constructed from cells cultured in a medium with bubbling air containing 5% carbon dioxide"	
BASE COUNT	114 a 157 c 153 g 78 t
ORIGIN	
Query Match	6.0%; Score 67.4; DB 10; Length 502;
Best Local Similarity	56.6%; Pred. No. 1.7e-09;
Matches	175; Conservative 0; Mismatches 116; Indels 18; Gaps 2;
BASE COUNT	112 a 150 c 147 g 73 t
ORIGIN	



Qy	218	--AACITCATCACACAGCAGGTTAAGTGCCTGGTGGCTGACATCCACTTCGACTATC	274	Db	337	CGGTGGCCATGATGGGGGG---ACGCCTTGAGAGATCCCGTGTGAACCCGGCACT	393
Db	286	ACCAGCTGTCAGGACAAGTACGAGCTGCACATCCACTTCGACTATC	345	Qy	335	TGGTAATG 343	
Qy	275	GCATGGCGCAAAGTAGCGAAATGGGCTGATGTCCTGGTATAACCTGGCAATA	334	Db	394	TGGCTACG 402	
Db	346	CGGGGCCATGATGGGGGG---ACGCCTTGAGAGATCCCGTGTGAACCCGGCACT	402				
Qy	335	TGGTAATG 343					
Db	403	TGGCTACG 411					
<b>RESULT 8</b>							
AV644087				LOCUS	AV626844		
DEFINITION	AV644087	Chlamydomonas reinhardtii 5'	mRNA sequence.	LOCUS	AV626844		
ACCESSION	AV644087			DEFINITION	AV626844		
VERSION	AV644087.1			ACCESSION	AV626844		
KEYWORDS	EST.			VERSION	AV626844.1		
SOURCE				KEYWORDS	EST.		
ORGANISM				ORGANISM	Chlamydomonas reinhardtii		
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;				Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;			
Chlamydomonadaceae; Chlamydomonas.				Chlamydomonadaceae; Chlamydomonas.			
I (bases 1 to 531)				REFERENCE	1 (bases 1 to 531)		
Asanizu, E., Miura, K., Kuro, K., Inoue, Y., Fukuzawa, H., Ohama, K., Nakamura, Y. and Tabata, S.				AUTHORS	Asanizu, E., Miura, K., Kuro, K., Inoue, Y., Fukuzawa, H., Ohama, K., Nakamura, Y. and Tabata, S.		
TITLE				TITLE			
Generation of expressed sequence tags from low-CO <sub>2</sub> and high-CO <sub>2</sub> adapted cells of Chlamydomonas reinhardtii				adapted cells of Chlamydomonas reinhardtii			
JOURNAL	DNA Res. 7 (5), 305-307 (2000)			JOURNAL	DNA Res. 7 (5), 305-307 (2000)		
MEDLINE	2053644			MEDLINE	2053644		
COMMENT				COMMENT			
The First Laboratory for Plant Gene Research				Contract: Erika Asanizu			
Kazusa DNA Research Institute				The First Laboratory for Plant Gene Research			
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan				Kazusa DNA Research Institute			
Email: asanizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.				Yana 1532-3, Kisarazu, Chiba 292-0812, Japan			
FEATURES				Email: asanizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.			
source				Location/Qualifiers			
1. .531				1. .545			
/organism="Chlamydomonas reinhardtii"				/organism="Chlamydomonas reinhardtii"			
/strain="C9"				/strain="C9"			
/db_xref="taxon:3055"				/db_xref="taxon:3055"			
/clone="HC081e11_r"				/clone="HC081e11_r"			
/clone_lib="Chlamydomonas reinhardtii SK;" Site:1: ECR1; Site:2:				/clone_lib="Chlamydomonas reinhardtii SK;" Site:1: ECR1; Site:2:			
xhoI: The cdna library was constructed from cells cultured in a medium with bubbling air containing 5% carbon dioxide"				xhoI: The cdna library was constructed from cells cultured in a carbon stress acclimatized condition in which carbon dioxide concentration in the bubbling gas was changed from 5% to 0.04%"			
BASE COUNT	120 a 168 c 162 g 81 t			BASE COUNT	121 a 177 c 166 g 81 t		
ORIGIN				ORIGIN			
Query Match	6.0%	Score 67.4; DB 10; Length 531;		Query Match	6.0%	Score 67.4; DB 10; Length 545;	
Best Local Similarity	56.6%	Pred. No. 1.8e-09;		Best Local Similarity	56.6%	Pred. No. 1.8e-09;	
Matches 175; Conservative	0	Mismatches 116; Indels 18; Gaps 2;		Matches 175; Conservative	0	Mismatches 116; Indels 18; Gaps 2;	
Query	50	TGGGAATGTCGGCATGGTGTCTCCATCGCGTAGCTCCATGACCAATAGC	109	Query	50	TGGGAATGTCGGCATGGTGTCTCCATCGCGTAGCTCCATGACCAATAGC	109
Best Local Similarity	56.6%	Pred. No. 1.8e-09;		Best Local Similarity	56.6%	Pred. No. 1.8e-09;	
Matches 175; Conservative	0	Mismatches 116; Indels 18; Gaps 2;		Matches 175; Conservative	0	Mismatches 116; Indels 18; Gaps 2;	
Query	97	TGGCAAGGGCGTGGCGACCCAGCGCATGCCTTCGACCATGACCAACGG	156	Query	90	TGGCAAGGGCGTGGCGACCCAGCGCATGCCTTCGACCATGACCAACGG	149
Best Local Similarity	56.6%	Pred. No. 1.8e-09;		Best Local Similarity	56.6%	Pred. No. 1.8e-09;	
Matches 175; Conservative	0	Mismatches 116; Indels 18; Gaps 2;		Matches 175; Conservative	0	Mismatches 116; Indels 18; Gaps 2;	
Query	110	GTCAGACAGACCGTGAAGCAGGTCATCAATCAAGCGCTGGAGGAGCTG	169	Query	110	GTCAGACAGACCGTGAAGCAGGTCATCAATCAAGCGCTGGAGGAGCTG	169
Best Local Similarity	56.6%	Pred. No. 1.8e-09;		Best Local Similarity	56.6%	Pred. No. 1.8e-09;	
Matches 175; Conservative	0	Mismatches 116; Indels 18; Gaps 2;		Matches 175; Conservative	0	Mismatches 116; Indels 18; Gaps 2;	
Query	157	ACACCGCGAATGTCGACGTGACGGTGGACCGAGCTGACCATGACCAACGG	216	Query	150	ACACCGCGAATGTCGACGTGACGGTGGACCGAGCTGACCATGACCAACGG	209
Best Local Similarity	56.6%	Pred. No. 1.8e-09;		Best Local Similarity	56.6%	Pred. No. 1.8e-09;	
Matches 175; Conservative	0	Mismatches 116; Indels 18; Gaps 2;		Matches 175; Conservative	0	Mismatches 116; Indels 18; Gaps 2;	
Query	170	ATATCGCTCGTGTATCGGACCGTGGACGGTGGACCGGGGAAGACGGTCA	217	Query	170	ATATCGCTCGTGTATCGGACCGTGGACGGTGGACCGGGGAAGACGGTCA	217
Best Local Similarity	56.6%	Pred. No. 1.8e-09;		Best Local Similarity	56.6%	Pred. No. 1.8e-09;	
Matches 175; Conservative	0	Mismatches 116; Indels 18; Gaps 2;		Matches 175; Conservative	0	Mismatches 116; Indels 18; Gaps 2;	
Query	217	ACATCGCTCGCATCACAGTGCAGGCAGAAGAGGGCTGGCATGAAGATTGCG	276	Query	210	ACATCGCTCGCATCACAGTGCAGGCAGAAGAGGGCTGGCATGAAGATTGCG	269
Best Local Similarity	56.6%	Pred. No. 1.8e-09;		Best Local Similarity	56.6%	Pred. No. 1.8e-09;	
Matches 175; Conservative	0	Mismatches 116; Indels 18; Gaps 2;		Matches 175; Conservative	0	Mismatches 116; Indels 18; Gaps 2;	
Query	218	--AACITCATCACACAGCGGTTAACGGTGCCTGGCTGACATCCACTTCGACTATC	274	Query	218	--AACITCATCACACAGCGGTTAACGGTGCCTGGCTGACATCCACTTCGACTATC	274
Best Local Similarity	56.6%	Pred. No. 1.8e-09;		Best Local Similarity	56.6%	Pred. No. 1.8e-09;	
Matches 175; Conservative	0	Mismatches 116; Indels 18; Gaps 2;		Matches 175; Conservative	0	Mismatches 116; Indels 18; Gaps 2;	
Query	277	ACGAGCTGTCAGGACAGTACGACCTGCGCTGGCGCCACATCCACTTCGACCA	336	Query	277	ACGAGCTGTCAGGACAGTACGACCTGCGCTGGCGCCACATCCACTTCGACCA	334
Best Local Similarity	56.6%	Pred. No. 1.8e-09;		Best Local Similarity	56.6%	Pred. No. 1.8e-09;	
Matches 175; Conservative	0	Mismatches 116; Indels 18; Gaps 2;		Matches 175; Conservative	0	Mismatches 116; Indels 18; Gaps 2;	
Query	275	GCATGGCGCTGAAAGTAGCGGATAACGGCGTGAATGCTGCTGATTAACCGCGCAATA	334	Query	275	GCATGGCGCTGAAAGTAGCGGATAACGGCGTGAATGCTGCTGATTAACCGCGCAATA	334
Best Local Similarity	56.6%	Pred. No. 1.8e-09;		Best Local Similarity	56.6%	Pred. No. 1.8e-09;	
Matches 175; Conservative	0	Mismatches 116; Indels 18; Gaps 2;		Matches 175; Conservative	0	Mismatches 116; Indels 18; Gaps 2;	
Query	335	TGGTAATG 343		Query	335	TGGTAATG 343	

RESULT	10	Db	387	TCGCTGACG	395
AV29427					
LOCUS	AV629427	467 bp	mRNA	linear	EST 15-DEC-2000
DEFINITION	reinhardtii. cDNA clone LC058f01_5'	5%	CO2	Chlamydomonas	reinhardtii 5% CO2 Chlamydomonas reinhardtii mRNA sequence.
ACCESSION	AV629427				
VERSION	AV629427.1				
KEYWORDS	EST.				
SOURCE	Chlamydomonas reinhardtii.				
ORGANISM	Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales; Chlamydomonadaceae; Chlamydomonas.				
REFERENCE	1 (bases 1 to 377)				
AUTHORS	Asamizu,E., Miura,K., Kuchio,K., Inoue,Y., Fukuzawa,H., Ohyama,K., Nakamura,Y. and Tabata,S.				
TITLE	Generation of expressed sequence tags from low-CO2 and high-CO2 adapted cells of Chlamydomonas reinhardtii				
JOURNAL	Kazusa DNA Res.	7 (5), 305-307 (2000)			
COMMENT	Contact: Erika Asamizu The First Laboratory for Plant Gene Research Kazusa DNA Research Institute Yana 1532-3, Kisarazu, Chiba 292-0012, Japan Email: asamizu@kazusa.or.jp, URL: <a href="http://www.kazusa.or.jp/en/plant/">http://www.kazusa.or.jp/en/plant/</a>				
FEATURES	source				
BASE COUNT	108	a	149	c	142
ORIGIN			g	t	68
Query Match	5.9%; Score 65.8; DB 10; Length 467; Best Local Similarity 56.3%; Pred. No. 5.1e-09; Matches 174; Conservative 0; Mismatches 117; Indels 18; Gaps 2;				
QY	50	TTCGGAAATGTGCCGATTTGGCGATGGTGTCTCCATGCCGTACAGTCATGACCAATACGC	109		
Db	93	TGGCAAGGTGCCGGGGCAGCACGCCATGCCCTCAGACCATGACCACCG	152		
QY	110	GTACGACAGACCTCGAAGCACCGTCAATCAAATCAAGGCCCTGGAAACCGGTGGCGCTG	169		
Db	153	ACACCCGAATGTGCACTGACCGTGACAGGTGAAGAAGTGTGCGGACGCCGGCGG	212		
QY	170	ATATGCCCGTGTATCCCTACCGACCATGGACGGGGAGAACGGTCA-----	217		
Db	213	ACATCGTCCGATCACAGTCAGGGCAGAGGAGCTGCGTCACTCCATT	272		
QY	218	--AACATCATCAAACGAGGTAAACGCGCCCTGTGCGTCACTCCACTGGACATATC	274		
Db	273	ACGAGCTGTCAAGGAGAAGTACGACCTGGCCGACATCCATT--CCAGC	329		
QY	275	GCATTGGCTGAAGTAGCGGATAACGGGTCGATGTCTCGTATTACCCCTGGCAATA	334		
Db	330	CCACGGGGGCCATGATGTGGCGGACGCCCTCGAGAAGATCCGTGTAACCCGGCAACT	389		
TCGGTAATG	343				
Db	390	TGCTGACG	398		
RESULT	12				
AT770308					
LOCUS	AT770308	380 bp	mRNA	linear	EST 24-JAN-2000
DEFINITION	35 Mycobacterium anaerobic stationary phase library				
ACCESSION	A1770308				
VERSION	A1770308.1				
KEYWORDS	EST.				
SOURCE	Mycobacterium smegmatis.				
ORGANISM	Mycobacterium smegmatis. Mycobacteriaceae; Actinomycetales; Bacteria; Actinobacteria; Actinomycetales; Mycobacterium.				
RESULT	11	AV641368	377 bp	mRNA	linear EST 15-DEC-2000
AV641368					
LOCUS	AV641368				
DEFINITION	Chlamydomonas reinhardtii 5% CO2 Chlamydomonas reinhardtii mRNA sequence.				
ACCESSION	AV641368				
VERSION	AV641368.1				
KEYWORDS	EST.				
SOURCE	Chlamydomonas reinhardtii.				
ORGANISM	Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales; Chlamydomonadaceae; Chlamydomonas.				
REFERENCE	1 (bases 1 to 377)				
AUTHORS	Asamizu,E., Miura,K., Kuchio,K., Inoue,Y., Fukuzawa,H., Ohyama,K., Nakamura,Y. and Tabata,S.				
TITLE	Generation of expressed sequence tags from low-CO2 and high-CO2 adapted cells of Chlamydomonas reinhardtii				
JOURNAL	Kazusa DNA Res.	7 (5), 305-307 (2000)			
COMMENT	Contact: Erika Asamizu The First Laboratory for Plant Gene Research Kazusa DNA Research Institute Yana 1532-3, Kisarazu, Chiba 292-0012, Japan Email: asamizu@kazusa.or.jp, URL: <a href="http://www.kazusa.or.jp/en/plant/">http://www.kazusa.or.jp/en/plant/</a>				
FEATURES	source				
BASE COUNT	89	a	118	c	116
ORIGIN			g	t	54
Query Match	5.5%; Score 62; DB 10; Length 377; Best Local Similarity 55.6%; Pred. No. 6.7e-08; Matches 150; Conservative 0; Mismatches 105; Indels 15; Gaps 1;				
QY	50	TTCGGAAATGTGCCGATTTGGCGATGGTGTCTCCATGCCGTACAGTCATGACCAATACGC	109		
Db	102	TGGCAAGGTGCCGGGGCAGCACGCCATGCCCTCAGACCATGACCACCG	161		
QY	110	GTACGACAGACCTCGAAGCACCGTCAATCAAATCAAGGCCCTGGAAACCGGTGGCGCTG	169		
Db	162	ACACCCGAATGTGCACTGACCGTGACAGGTGAAGAAGTGTGCGGACGCCGGCGG	221		
QY	170	ATATGCCCGTGTATCCCTACCGACCATGGACGGGGAGAACGGTCA-----	217		
Db	222	ACATCGTCCGATCACAGTCAGGGCAGAGGAGGAGGTGAGCGCTGATCGCG	281		
QY	218	--AACATCATCAAACGAGGTAAACGCGCCCTGTGCGTCACTCCACTGGACATATC	274		
Db	282	AGCAGCTGTCAAGGAGAAGTACGACGTGGCCCTGTGGCCGACATCCACTTCCAGCCA	341		
QY	275	GCATTGGCTGAAGTAGCGGATAACGGGTCGATGTCTCGTATTACCCCTGGCAATA	304		
Db	342	CGTGGCCATGATGTGCGGGAGGCCCTCG	371		

REFERENCE	1 (bases 1 to 380)	Chlamydomonadaceae; Chlamydomonas.
AUTHORS	Murugasu-Oei,B., Tay,A. and Dick,T.	
TITLE	Upregulation of stress response genes and ABC transporters in anaerobic stationary-phase <i>Mycobacterium smegmatis</i> .	
JOURNAL	<i>Mol. Gen. Genet.</i> 262 (4-5), 677-682 (1999)	
COMMENT	Contact: Murugasu-Oei, B. <i>Mycobacterium Laboratory</i> Institute of Molecular and Cell Biology 30 Medical Drive, Singapore 117609, Republic of Singapore	
MEDLINE	Tel: 65 874 3011 Fax: 65 779 1117 Email: mbobone@imcb.nus.edu.sg Insert Length: 380 Std Error: 0.00	
FEATURES	source	
	Location/Qualifiers	
	1. .380	
	/organism="Mycobacterium smegmatis"	
	/strain="mc2155"	
	/db_xref="taxon:1772"	
	/clone_1lib="Mycobacterium anaerobic stationary phase library"	
	/lab_host="E. coli XL1-Blue MRF"	
	/note="Vector: Lambda ZAP II; Bacilli were disrupted using a Mini bead beater (Biospec). RNA was isolated using the RNeasy protocol (QIAGEN). Purified RNA was subjected to 2 rounds of digestion with RNase-free DNase I (Promega). DNase I was heat-inactivated at 75°C for 5 min. and removed by using RNeasy columns followed by phenol extraction and ethanol precipitation. The RNA preparations were confirmed to be free of genomic DNA contamination by carrying out PCR and RT-PCR using the Access kit (Promega) and primers specific for the histone-like protein gene hlp (Lee et al., 1998). cDNA was synthesized using random hexamer primers (Promega) and Stratagene's cDNA synthesis kit. cDNA fragments were ligated into Lambda ZAP II vector and packaged in vitro using Stratagene reagents. Vector and packaged in vitro	
BASE COUNT	68 a 130 c 121 g 61 t	
ORIGIN		
Query Match	5.0%: Score 56.4; DB 9;	
Best Local Similarity	53.7%; Pred. No. 3.7e-05;	
Matches	117; Conservative 0; Mismatches 101; Indels 0; Gaps 0;	
QY	47 AGTTGGAACTGTGCCATTGCGATGTCGCCTCCATGACGATTCAGTACCAATA 106	
Db	163 ATGTGGCGCCATGGCATCGCAGCACCCGATCCAGTCACAGTACCA 222	
QY	107 CGGTACGACAGACGTGCAAGAACGGCTCAATCAATCAAGGGCTTGACCGT 166	
Db	223 CCAAGACCACAGTCACACTCACAGTCACAGTCAGCTCGGGT 282	
QY	167 CGATATCGTCGTGTAACCGTACCGACGATGACGGGGCAAGGCTTCAACTCATCA 226	
Db	283 GTGACATCGTGGTGGCGCGCTGAGGAGACGCCAGCCTGGCGAGATCG 342	
QY	227 AACAGCAGGTTAACGGCGCGCTGAGTACATCC 264	
Db	343 CGGGCACAGCAGATGCCGGTACGCCGATCCAC 380	
RESULT 13		
AV640429	AV640429	303 bp mRNA linear EST 15-DEC-2000
DEFINITION	AV640429 chlamydomonas reinhardtii 5' CO2	Chlamydomonas reinhardtii
ACCESSION	AV640429	cDNA clone HCL016b07_r 5', mRNA sequence.
VERSION	AV640429.1	GI:10783757
KEYWORDS	EST.	
SOURCE	Chlamydomonas reinhardtii	
ORGANISM	Chlamydomonas reinhardtii	
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoidae; Andropogoneae; Zea.		
REFERENCE	1 (bases 1 to 597).	
AUTHORS	Walbot,V.	
JOURNAL	Maize ESTs from various cDNA libraries sequenced at Stanford University Unpublished (1999)	
COMMENT	Contact: Walbot,V Department of Biological Sciences Stanford University 855 California Ave, Palo Alto, CA 94304, USA Tel: 650 723 2227 Fax: 650 725 8221 Email: walbot@stanford.edu	

FEATURES	Plate: 829011 row: H column: 05.
SOURCE	<p>/clone.lib="Stevia field grown leaf cDNA"</p> <p>/tissue.type="leaf"</p> <p>/dev_stage="field grown, mid-size"</p> <p>/lab_host="E. coli strain XLOLR"</p> <p>/note="vector: pRK-CMV; site_1: EcoRI; site_2: XbaI; this cDNA library was constructed from polyA+ enriched mRNA from field grown leaves. Mid-size actively growing leaves were collected and pooled from several plants and frozen immediately after harvesting in liquid nitrogen. The cDNA was prepared using an XbaI Poly(dT) linker-primer. An EcoRI adapter was ligated to the blunt end cDNA and the products were digested with EcoRI and XbaI enabling directional cloning into the Lambda ZAP Express vector. The library was amplified using the host strain XLI-Blue MRF+. Mass excision of the library was performed to obtain PBK-CMV phagemid clones in the host strain XLOLR. Single pass DNA sequencing was performed using the T3 promoter primer: 5' ATTAACCTCACTAAAGGGA 3'. This library was constructed by Alex Richman."</p>
BASE COUNT	157 a 139 c 187 g 113 t 1 others
ORIGIN	Query Match 4.8%; Score 53.8; DB 10; Length 597; Best local Similarity 54.0%; Pred. No. 3.e-05; Matches 109; Conservative 0; Mismatches 93; Indels 0; Gaps 0;
QY	26 GTGAAATTCACAGCTATTAGTTGGATGTGCCATTGGATGGTGCGCCATCG 85
Db	393 GGAGGAACACTCGAACCTGGATGTGGATGATCATGCCATAA 452
QY	86 CCCAACAGTCAGGACCAATACCGTAGACACAGCTGAAAGAACGGTCAAATCA 145
Db	453 GGATTCACACATGACGACTTCGATACAAAGATGTCGAAACAAACAGTAGAGGAGGTA 512
QY	146 AGCGCTGGAACCGTTGGCGATATGTCGTTACCGAACGGATGGACGGG 205
Db	513 TGAGGATACAGATAAAGAGCTGATCTGNTAGATAACAGTCAGGGTAGGAAGGAG 572
QY	206 CAGAACGGTCAACTCACTCAA 227
Db	573 CTGATGCTGCTTGAGATCAA 594
RESULT 15	
LOCUS	BG525868
DEFINITION	5'-3'-1-2 r Stevia field grown leaf cDNA Stevia rebaudiana cDNA 5', mRNA sequence.
ACCESSION	BG525868
VERSION	BG525868.1 GI:16949329
KEYWORDS	EST.
SOURCE	Stevia rebaudiana.
ORGANISM	Bukaryota; Viridiplantae; Streptophyta; Tracheophyta; Spermatophytina; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; eusterids II; Asterales; Asteraceae; Asteroideae; Heliantheae; Stevia.
REFERENCE	1 (bases 1 to 728)
AUTHORS	Brandle, J.E., Richman, A., Swanson, A.K. and Chapman, B.P.
TITLE	Leaf ESTs from Stevia rebaudiana: A resource for gene discovery in diterpene synthesis
JOURNAL	Unpublished (2001)
COMMENT	Contact: Jim Brandle Genomics and Biotechnology Agriculture and Agri-Food Canada - SCFPRC 1391 Sandford St., London, Ontario, CANADA, N5V 4T3 Tel: 519 457 1470 Fax: 519 457 3997 Email: brandle@em.agr.ca Seq primer: T3 promoter primer. Location/Qualifiers
FEATURES	<p>Search completed: May 10, 2003, 16:41:00</p> <p>Job time : 1006 secs</p>
SOURCE	<p>/organism="Stevia rebaudiana"</p> <p>/strain="751/1501"</p> <p>/cultivar="Lndrace"</p>

Gencore version 5.1.4\_ps-4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model.

Run on: May 10, 2003, 15:09:17 ; Search time 78 Seconds  
(without alignments)  
4399.632 Million cell updates/sec

Title: US-09-921-992-3  
Perfect score: 1119  
Sequence: 1 atgcataaccaggctcaat.....ttcaagcagggtgaaaaataa 1119  
Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 892724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
listing first 45 summaries

Database : Issued\_Patents\_NA:\*

1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:/\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:/\*  
3: /cgn2\_6/ptodata/1/ina/5C\_COMB.seq:/\*  
4: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:/\*  
5: /cgn2\_6/ptodata/1/ina/pctus\_COMB.seq:/\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:/\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

**SUMMARIES**

Result No.	Score	Query Match Length	DB ID	Description
1	1010	90.3	1010	2 US-09-827-190-8
2	1010	90.3	1010	4 US-09-170-187-8
3	631.6	50.4	1345	2 US-09-827-190-4
4	631.6	50.4	1345	4 US-09-170-187-4
5	619.8	55.4	1079	2 US-09-827-190-7
6	619.8	55.4	1079	4 US-09-170-187-7
7	534.6	47.8	886	2 US-09-827-190-10
8	534.6	47.8	886	4 US-09-170-187-10
9	269	24.0	886	2 US-09-827-190-9
10	55.4	5.0	1083	4 US-09-221-017B-309
11	55.4	5.0	1083	4 US-09-221-017B-309
12	40.6	3.6	1206	4 US-09-153-599A-2
13	40.6	3.6	2299	4 US-09-153-599A-1
14	38.4	3.4	1665	4 US-09-221-017B-1017
15	38.2	3.4	4411529	4 US-09-103-840A-1
16	35.4	3.2	583	4 US-09-998-416-1039
17	35.4	3.2	737	4 US-09-998-416-1039
18	35	3.1	1383	3 US-09-618-651A-7
19	35	3.1	1383	4 US-09-215-252-6
20	35	3.1	1383	4 US-09-215-252-6
21	34	3.0	8910	3 US-09-779-764A-1
22	34	3.0	4411529	4 US-09-103-840A-1
23	33.8	3.0	44377	2 US-09-804-227C-7
24	33.8	3.0	44377	2 US-09-804-198-1
25	33.6	3.0	606	2 US-09-242-871C-48
26	33.6	3.0	606	4 US-09-270-956-48
27	33.4	3.0	1660	4 US-09-297-468-1

**ALIGNMENTS**

RESULT 1-US-08-827-190-8

; Sequence 8, Application US/08827190  
; Patent No. 5858367

; GENERAL INFORMATION:

; APPLICANT: Rather, Philip N

; TITLE OF INVENTION: Methods For Screening For Antimicrobials  
; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Medlen & Carroll, LLP

; STREET: 220 Montgomery Street, Suite 2200

; CITY: San Francisco

; STATE: California

; COUNTRY: United States of America

; ZIP: 94104

; COMPUTER READABLE FORM:

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/827,190

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Carroll, Peter G.

; REGISTRATION NUMBER: 32,837

; REFERENCE/DOCKET NUMBER: CASE-02443

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 705-8410

; TELEFAX: (415) 397-8338

; INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1010 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: Unknown

; MOLECULE TYPE: DNA (genomic)

US-08-827-190-8

Query Match Best Local Similarity 90.3%; Score 1010; DB 2; Length 1010; Matches 1010; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Sequence 1, Appli Qy 1 ATGCATACAGGCTCCATTCAACTAGAAATCACGCTTACCTTGGAATGCG 60

Sequence 2, Appli Sequence 1, Appli Db 1 ATGCATACAGGCTCCATTCAACTAGAAATCACGCTTACCTTGGAATGCG 60

Sequence 3, Appli Sequence 4, Appli Sequence 5, Appli Sequence 6, Appli Sequence 7, Appli Sequence 8, Appli Sequence 9, Appli Sequence 10, Appli Sequence 11, Appli Sequence 12, Appli Sequence 13, Appli Sequence 14, Appli Sequence 15, Appli Sequence 16, Appli Sequence 17, Appli Sequence 18, Appli Sequence 19, Appli Sequence 20, Appli Sequence 21, Appli Sequence 22, Appli Sequence 23, Appli Sequence 24, Appli Sequence 25, Appli Sequence 26, Appli Sequence 27, Appli Sequence 28, Appli Sequence 29, Appli Sequence 30, Appli Sequence 31, Appli Sequence 32, Appli Sequence 33, Appli Sequence 34, Appli Sequence 35, Appli Sequence 36, Appli Sequence 37, Appli Sequence 38, Appli Sequence 39, Appli Sequence 40, Appli Sequence 41, Appli Sequence 42, Appli Sequence 43, Appli Sequence 44, Appli Sequence 45, Appli Sequence 46, Appli Sequence 47, Appli Sequence 48, Appli Sequence 49, Appli Sequence 50, Appli Sequence 51, Appli Sequence 52, Appli Sequence 53, Appli Sequence 54, Appli Sequence 55, Appli Sequence 56, Appli Sequence 57, Appli Sequence 58, Appli Sequence 59, Appli Sequence 60, Appli Sequence 61, Appli Sequence 62, Appli Sequence 63, Appli Sequence 64, Appli Sequence 65, Appli Sequence 66, Appli Sequence 67, Appli Sequence 68, Appli Sequence 69, Appli Sequence 70, Appli Sequence 71, Appli Sequence 72, Appli Sequence 73, Appli Sequence 74, Appli Sequence 75, Appli Sequence 76, Appli Sequence 77, Appli Sequence 78, Appli Sequence 79, Appli Sequence 80, Appli Sequence 81, Appli Sequence 82, Appli Sequence 83, Appli Sequence 84, Appli Sequence 85, Appli Sequence 86, Appli Sequence 87, Appli Sequence 88, Appli Sequence 89, Appli Sequence 90, Appli Sequence 91, Appli Sequence 92, Appli Sequence 93, Appli Sequence 94, Appli Sequence 95, Appli Sequence 96, Appli Sequence 97, Appli Sequence 98, Appli Sequence 99, Appli Sequence 100, Appli Sequence 101, Appli Sequence 102, Appli Sequence 103, Appli Sequence 104, Appli Sequence 105, Appli Sequence 106, Appli Sequence 107, Appli Sequence 108, Appli Sequence 109, Appli Sequence 110, Appli Sequence 111, Appli Sequence 112, Appli Sequence 113, Appli Sequence 114, Appli Sequence 115, Appli Sequence 116, Appli Sequence 117, Appli Sequence 118, Appli Sequence 119, Appli Sequence 120, Appli



481 GCGCATGGCTGTCATGATCACTCGATCGCAGTCAGCGTGT 540

QY 541 AAAGCCTCTGAGTCCTCCCTGCTGTGTTGAGCTTATCGTTGCTGGCAAACAGATCGAT 600

Db 541 AAAGCCTCTGAGTCCTCCCTGCTGTGTTGAGCTTATCGTTGCTGGCAAACAGATCGAT 600

QY 601 CAGCGTTGAGTACTGGGAACACCGAAGAGCTGGCTGCGCAGGGGAGATAATCC 660

Db 601 CAGCGTTGAGTACTGGGAACACCGAAGAGCTGGCTGCGCAGGGGAGATAATCC 660

QY 661 GGCATTTGGTTAGGCTGCTGCTGCTGAAAGATCAACCTGGACACCGTGGCTATCGCTG 720

Db 661 GGCATTTGGTTAGGCTGCTGCTGCTGAAAGATCAACCTGGACACCGTGGCTATCGCTG 720

QY 661 GGCATTTGGTTAGGCTGCTGCTGAAAGATCAACCTGGACACCGTGGCTATCGCTG 720

Db 661 GGCATTTGGTTAGGCTGCTGCTGAAAGATCAACCTGGACACCGTGGCTATCGCTG 720

QY 721 CGGGCGATCGGCTGGAAAGATCAAAAGTGATTTGAAATGCTGGTAC 780

Db 721 CGGGCGATCGGCTGGAAAGATCAAAAGTGATTTGAAATGCTGGTAC 780

QY 721 CGGGCGATCGGCTGGAAAGATCAAAAGTGATTTGAAATGCTGGTAC 780

Db 721 CGGGCGATCGGCTGGAAAGATCAAAAGTGATTTGAAATGCTGGTAC 780

QY 781 CGTTGGCAGAGGATCACTTCATCGCTGCCGACCTGTTGCGTCAGGATTATGTT 840

Db 781 CGTTGGCAGAGGATCACTTCATCGCTGCCGACCTGTTGCGTCAGGATTATGTT 840

QY 781 CGTTGGCAGAGGATCACTTCATCGCTGCCGACCTGTTGCGTCAGGATTATGTT 840

Db 781 CGTTGGCAGAGGATCACTTCATCGCTGCCGACCTGTTGCGTCAGGATTATGTT 840

QY 841 ATCGGTACGGTTAAGGGCTGAGGACACGGCTGGAGATTCATCACTCGATGGCT 900

Db 841 ATCGGTACGGTTAAGGGCTGAGGACACGGCTGGAGATTCATCACTCGATGGCT 900

QY 841 ATCGGTACGGTTAAGGGCTGAGGACACGGCTGGAGATTCATCACTCGATGGCT 900

Db 841 ATCGGTACGGTTAAGGGCTGAGGACACGGCTGGAGATTCATCACTCGATGGCT 900

QY 901 TCGATTTATGGCTGGCTGGGAATGGCCCGAGGTGAGGCTGGTTCTAGACTCGCTC 960

Db 901 TCGATTTATGGCTGGCTGGGAATGGCCCGAGGTGAGGCTGGTTCTAGACTCGCTC 960

QY 901 TCGATTTATGGCTGGCTGGGAATGGCCCGAGGTGAGGCTGGTTCTAGACTCGCTC 960

Db 901 TCGATTTATGGCTGGCTGGGAATGGCCCGAGGTGAGGCTGGTTCTAGACTCGCTC 960

QY 961 ACCGGGGCACAAAGAAAACGGCCCTATGAGATGGCCGGCAAGA 1010

Db 961 ACCGGGGCACAAAGAAAACGGCCCTATGAGATGGCCGGCAAGA 1010

QY 961 ACCGGGGCACAAAGAAAACGGCCCTATGAGATGGCCGGCAAGA 1010

Db 961 ACCGGGGCACAAAGAAAACGGCCCTATGAGATGGCCGGCAAGA 1010

RESULT 3  
US-08-827-190-4

Sequence 4, Application US/08827190

PATENT NO. 5858367

GENERAL INFORMATION:

APPLICANT: Rather, Philip N.

TITLE OF INVENTION: Methods For Screening For Antimicrobials Utilizing aarc And Compositions Thereof

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Madlen & Carroll, LLP  
STREET: 220 Montgomery Street, Suite 2200  
CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATOR: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/827,190

FILING DATE:

CLASSIFICATION: 435 ATTORNEY/AGENT INFORMATION:

NAME: Carroll, Peter G.  
REGISTRATION NUMBER: 32,837  
REFERENCE DOCKET NUMBER: CASE-02443

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 705-8110  
TELEFAX: (415) 397-8338

INFORMATION FOR SSO ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1345 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear

FEATURE: NAME/KEY: CDS  
LOCATION: 175..1272  
OTHER INFORMATION: /gene= "aarC"  
OTHER INFORMATION: /note= "Similar to E. coli Gepe protein listed by GenBank"

Query Match  
Best Local Similarity 74.5%; Score 631.6; DB 2; Length 1345;  
Matches 809; Conservative 0; Mismatches 274; Indels 3; Gaps 1;

us-08-827-190-4

QY 1 ATGCATTAACCGGGCTCAATCAACCTGAGAATTCACAGCTTACGTTAGGATTGATG 56.4%; Score 631.6; DB 2; Length 1345;  
Db 175 ATGCATTAATGATCACCGATAAAAGACGTAATTCACCCGAAATTATGTA3GTAACTG 56.4%; Score 631.6; DB 2; Length 1345;  
Db 61 CGCATTTGGCGATGCTCCATCAACCTGAGAATTCACCCGAAATTATGTA3GTAACTG 56.4%; Score 631.6; DB 2; Length 1345;  
Db 235 CCTATTCGGATGGTCTCCATCGCTGAGGATTCACCTCGCTGAGGATTCACCTCGCTG 56.4%; Score 631.6; DB 2; Length 1345;  
Db 121 GTCGAAGCAAGGTCATCAATCAAGGCGCTGGAAGACGCGTTGCGCTGATATCGTCCGT 56.4%; Score 631.6; DB 2; Length 1345;  
Db 295 GTTGAAGCCACTGTGGCAATTCACATCACTTGACGCTGTTAGGTTGATATCGGCCG 56.4%; Score 631.6; DB 2; Length 1345;  
Db 181 GTATCGTACCGACGATGGACGGGGAGAACGCGTCAACTCATCAACACAGGTTAAC 56.4%; Score 631.6; DB 2; Length 1345;  
Db 355 GTGTCCTGTTCTACAGTGGATGCTGAGGAGACGCGCTTAAATTAAATAGCAGCGCTGAAT 56.4%; Score 631.6; DB 2; Length 1345;  
Db 241 GTGCCCTGGGGCTGACATCACTTGACGCTGACTATCGCATTCGCTGAGTGGAAAGTGGAAATAC 56.4%; Score 631.6; DB 2; Length 1345;  
Db 415 GNGCCATTTGGTGGCAATTCACCTTGACTTGTGAACTGATGGCTGAGGAGGTGGTGAAT 56.4%; Score 631.6; DB 2; Length 1345;  
Db 361 GNGGTGACTGTGGCTGATGGCTAACACATTGGATCGTATGGTTAATGAGAGCGTATTCG 56.4%; Score 631.6; DB 2; Length 1345;  
Db 535 GTCGTGATAGTGGCTCATCACACAGATTCCTATCGTATAGGGCTAATGGCGGTCA 56.4%; Score 631.6; DB 2; Length 1345;  
Db 421 CTGGAAGAAGTCTGAGAGAAAGTITGGCAGACCGACCGCGCAGCGCTGCTGAGATCT 56.4%; Score 631.6; DB 2; Length 1345;  
Db 595 CTGGAAGAAGTATCCAGAAAAATGGCTGAGCCACACTGGATCATGGTATCCAA 56.4%; Score 631.6; DB 2; Length 1345;  
Db 481 GGCATGGTCATGATCATOTCGATGCCGCTGAGACTTCGATCGTCAMGTCAAGCGTG 56.4%; Score 631.6; DB 2; Length 1345;  
Db 655 GCAATGCGACATGTTGATATTTGGACGGGTGAATTCGATCAGCTCAAGCTCAGTGT 56.4%; Score 631.6; DB 2; Length 1345;  
Db 541 AAAGCCGCTGCTCCCTCGCTGTTGAGCTTATCGTTGCTGSCAAACAGATCGAT 56.4%; Score 631.6; DB 2; Length 1345;  
Db 715 AAGCTGTTGGATGTCCTCTGGCGGGCTTATCGTTATGGCCAAAGATGAT 56.4%; Score 631.6; DB 2; Length 1345;  
Qy 601 CAGCCGTTGCACTGGGATCACCGAGGCCGGGGCGCAGGGGCGTAATATCC 660  
Db 601 CAGCCGTTGCACTGGGATCACCGAGGCCGGGGCGCAGGGGCGTAATATCC 660  
Db 775 CAACCCCTCTACCTCGTCAATTCAGAAGGCTGGGGCTCTCGTCACTGTAATCA 834  
Db 661 GGCATTTGGTTAGGCTGCTGCTGCTGAGGCTGGCGACACGGTGGCTATCGCTG 720  
Db 835 GCAATGCGTCATGTTGATGTTGTTGAGGATGTTGGCTGAGGATTCACGTTACCTC 894  
Qy 721 CGGGCGATCGGCTGGAAAGATCAAAAGTGATTTGAAATGCTGGCTATCGCTG 780  
Db 895 GGGCAGATCCCTGGAGGAAGTGAAGTCGGCTGGCTGAGGATTCACGCTACG 954  
Qy 781 CGTTCGAGGATCACTTCATCGCTGCCGACCTGTTGGCTGAGGATTTGATGTT 840  
Db 955 CCTCGTGGCATCTTGCTGAGGATTCACCTTGCTGCGCCACCGCTGAGGATTCACGCTACG 1014  
Qy 841 ATCGGTACGGTTAACGGCTGGAGGACACGCCCTGGAGAATTCACCTCCGATGGCGT 900  
Db 1015 ATGGGACGGTAATCTTGGAGCAGGCCCTGAGAATTCACGCCATGGATGTC 1074  
Qy 901 TCGATTTGGCTGGTGAATGGCCGCCAGGGTGAAGCGCTGTTCTACACTCGCGTC 960

**RESULT 4**  
**US-09-170-187-4**  
**Sequence 4, Application US/09/170187**  
**Patent No. 6383745**

**GENERAL INFORMATION:**

**APPLICANT:** Rather, Philip N.  
**TITLE OF INVENTION:** Methods For Screening For Antimicrobials  
**NUMBER OF SEQUENCES:** 13  
**CORRESPONDENCE ADDRESS:**  
**ADDRESSEE:** Medien & Carroll, LLP  
**STREET:** 220 Montgomery Street, Suite 2200  
**STATE:** California  
**COUNTRY:** United States of America  
**ZIP:** 94104

**COMPUTER READABLE FORM:**

**MEDIUM TYPE:** Floppy disk  
**COMPUTER:** IBM PC compatible  
**OPERATING SYSTEM:** PC DOS/MS-DOS

**SOFTWARE:** PatentIn Release #1.0, Version #1.30

**CURRENT APPLICATION DATA:**  
**APPLICATION NUMBER:** US/09/170,187  
**FILING DATE:**  
**CLASSIFICATION:**

**PRIOR APPLICATION DATA:**  
**APPLICATION NUMBER:** 08/627,190  
**FILING DATE:**  
**ATTORNEY/AGENT INFORMATION:**  
**NAME:** Carroll, Peter G.  
**REGISTRATION NUMBER:** 32,837  
**REFERENCE/DOCKET NUMBER:** CASE-02443

**TELECOMMUNICATION INFORMATION:**  
**TELEPHONE:** (415) 705-8110  
**TELEFAX:** (415) 397-8388  
**INFORMATION FOR SEQ ID NO: 4:**  
**SEQUENCE CHARACTERISTICS:**  
**LENGTH:** 1345 base pairs  
**TYPE:** nucleic acid  
**STRANDEDNESS:** double  
**TOPOLOGY:** linear  
**MOLECULE TYPE:** DNA (genomic)  
**FEATURE:**  
**NAME/KEY:** CDS  
**LOCATION:** 175..1272  
**OTHER INFORMATION:** /gene= "aarc"  
**OTHER INFORMATION:** /note= "Similar to E. coli Gcpe protein listed by GenBank."  
**US-09-170-187-4**

**Query Match** 56..48; Score 631..6; DB 4; length 1345;  
**Best Local Similarity** 74..53; Pred. No. 2e-189; Indels 3; Gaps 1;  
**Matches** 809; Conservative

**QY** 1 ATGCATACCGGCTCAATCAGCTGAAATCACACGTTAACGTGTTAGTGGAAATGTG 60  
**Db** 175 ATGCATATGATCACCGATAAAAGCTGTAATCCACCGGAATTATGTAACGTG 234

**QY** 61 CCATTTGGGATGGTCCTTCAATCGGCCATACGTCATGACCAATAGGGTAGACAGAC 120  
**Db** 235 CCTATTGGGATGGTCCTCATGTCCTCAACTATGACGATACGGCACAGGGAT 294  
**QY** 121 GTGAAAGACGGTCAATCAATCAAGGGCTGAAAGGCTGGCTGATATCGTCGT 180  
**Db** 295 GTGAAAGACGGTCAATCAATCAAGGGCTGAAAGGCTGGCTGATATCGTCGT 354  
**QY** 181 GTATCCGTTACCGACGATGGACGGGAGAGAGCTCAAACTCATCAAAGAGGTTAAC 240  
**Db** 355 GTGCTGTCTTACATGATGATGAGCAGAGCCTTAATTAATGAGCCGCGTGAAT 414  
**QY** 241 GTGCCGCGTGGCTGACTGACATCCACTTGACTATGCAATGCGATGCTGAAAGTACGGGAAATAC 300  
**Db** 415 GTGCTGTCTTACATGATGATGAGCAGAGCCTTAATTAATGAGCCGCGTGAAT 474  
**QY** 301 GGGTCGATGTCGCTTAAACCTGGATAATCGGATATCGGATAGAGGGTATTCGAT 360  
**Db** 475 GGTTGGTGAATGCTACGAAATTAAACCCAGGTAAATCGGAGTGGCTGAAAGGCTGATTCGCA 534  
**QY** 361 GTGCTGTACTGCTCCTGACATCCAGTCAACATTCCTGAGTGGCTGAAACCCGGATC 594  
**Db** 535 GTCTGTTGATAGTCCTGCTACACATTCCTGATAC 594

**QY** 421 CTGGAAAAGAGCTGCAAGAAAGTAGGGGACCCGGCCAGGGTTGGAAATCT 480  
**Db** 595 CTGGAAAAGAGATATCCAGAAATAACCGGAACTGGGACCCAAACACTGAGATGGTGAATCA 654  
**QY** 481 GGCATGGCTCATGTGATCATCGATGCCGTAACCTGATCAGTCAGTCAGTCAGCTG 540  
**Db** 655 GCATGGCATGTGATCTGGGATGAGCTGCAATTCGAGCTGATTCAGTCAGTGT 714  
**QY** 541 AAAGCGTGTGACGCTCTCCCTCGCTGTTGAGCTTATGCTTGTGCAAAACAGATCAT 600  
**Db** 715 AAAGCGTGTGAGCTCTTCTTGTGCGTGGCTTATGCTTAAATGGGGCAAAATGAT 774  
**QY** 661 GGCATGGCTCATGGGATCACCGAAGCGGTTGGCTGAGCGAGGAGCAATTC 660  
**Db** 775 CAACCACTTACCTCGGTTAACAGAAGGGTGGGCTCTGGTTCTGGTTCTGGTT 834  
**QY** 662 CGCCATGGCTTGGGATCACCGAAGCGGTTGGCTGAGCGAGGAGCAATTC 720  
**Db** 835 GCATTTGGCTTGGTGTGCTGGCTGAGGATATCGGGATAGTGGCTGAGTAC 894  
**QY** 721 GCGGCCGATCCGGCTGCAAGAGATCAAGTGGGTTGATATTGAAATCGTGGPATC 780  
**Db** 895 GCGCAGATCTCTTGTGAGGAATGAGATGGCTTATGCTTAAATGCTTACGGATC 954  
**QY** 781 CGTTGGCTGGGATCACTCATGCCGCTGGGACCTTGCGTGGCTGAGGATTC 840  
**Db** 955 CGCTCACGGCATCACTTATGCTGAGGATATCGGCCAACCTGTCAGCCAAAGT 1014  
**QY** 841 ATGGTAGGTTAACGCCCTGGAGCAACGCCCTGAGAATATCATCACCCGAGCTG 900  
**Db** 1015 ATGGTAGGTTAACGCCCTGGAGCAACGCCCTGAGAATATCATCACCCGAGCTG 1074  
**QY** 901 TCGATTATGGCTGGCTGAGGATGGCCAGGGAGGCTGCTGTTACACTCGGGTC 960  
**Db** 1075 TCGATTATGGCTGGCTGAGGATGGCCCTGGAGGATGGCTGCTGTTACACTTG 1134  
**QY** 961 ACGGGGGACGAGAGAAAGCCCTATGAGATGCCCTGAGAATATCATCACCCGAGCTG 1017  
**Db** 1135 GTGGCCGCAACCAAAGTGGTTATGAGATGCCCTGGAGGATGGCTGCTGTT 1194  
**QY** 1018 GACACACAGATATGATGCCCTGGAGAAGCCAGTGGCTGAGGATGGCTGCTG 1077  
**Db** 1195 GATATGACATTTATGATGCCCTGGAGAATTCGCCAAAGACACAACTT 1254  
**QY** 1078 GACGAA 1083  
**Db** 1255 GATGAA 1260

RESULT 5  
 US-08-827-190-7  
 ; Sequence 7, Application US/08827190  
 ; Patent No. 5558367  
 GENERAL INFORMATION:  
 ; APPLICANT: Rather, Philip N.  
 TITLE OF INVENTION: Methods For Screening For Antimicrobials  
 TITLE OF INVENTION: Utilizing aarc And Compositions Thereof  
 NUMBER OF SEQUENCES: 13  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Medlen & Carroll, LLP  
 STREET: 220 Montgomery Street, Suite 2200  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: United States of America  
 ZIP: 94104  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/827,190  
 FILING DATE:  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Carroll, Peter G.  
 REGISTRATION NUMBER: 32,837  
 REFERENCE/DOCKET NUMBER: CASE-02443  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 705-8410  
 TELEFAX: (415) 397-8338  
 INFORMATION FOR SEQ ID NO: 7:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1079 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: unknown  
 MOLECULE TYPE: DNA (genomic)  
 US-08-827-190-7

Query Match 55.4%; Score 619.8; DB 2; Length 1079;  
 Best Local Similarity 74.9%; Pred. No. 9.4e-186; Indels 3; Gaps 1;  
 Matches 790; Conservative 0; Mismatches 262;

Qy 17 CATTGCAACGAGAAATCAACAGCSTATTAGTGGAAATGGCGATTCGGATGGTG 76  
 Db 4 CTATTAGCGCTGTGATCGACCAAATTATGGGGAAATGTTGGGATG 63

Qy 77 CTCCCATGCCATCAGTCATGACCAATGCCGTACGACAGACGCTGACCAACGCTCA 136  
 Db 64 CCCTTAATGCGCTGTGATCGACCAAATTATGGGGAAATGTTGGGATG 123

Qy 137 ATCAAATCAAGGCCCTGGAAGCGGTGGCGCTGATTCGGTATCCCTACCGACGA 196  
 Db 124 CTCAAATTAACATTAGAAGCTGCTGAGATATTGTTGATCTGCACAA 183

Qy 197 TGGACSGGGCAGAACGCTTCAAACTCATCAACAGCAGGTAAACGCCGGTGGCTG 256  
 Db 184 TGGATGCTGCCGAGATTAACAAATTAAACAGAGTAGTTCGGCTCGTAC 243

Qy 257 ACATCCACTTGCCTACTATGCATGGCTGAAGTAGCGGAATACCGGGTCTGATGGCTGC 316  
 Db 244 ATATTCATTCGCACTATCGTATCGCTTAAGCTCGAGAAATATGGAGTGTGTTAC 303

Qy 317 GTTAATACCGCGCAATATCGCTAATGAGAGCGGTATTCGAACTGGGTGACTGTCGC 376  
 Db 304 GATCAATTCGCACTATGGCTGCAACATTGGCTGCGATGGCTGCGTGGCTGCGC 363

Qy 377 GGATAAACATTCGATCCGATCCGATTCACGGGATACGCCGATCGCTGAGAAAGATCTGC 436  
 Db 364 GAGACAAATTCGATTCGATTCGATTCGATTCGATTCGAGGCTCTTGAAGAAAGATTG 423

Query Match 54.3%; Score 437.0; DB 1; Length 1058  
 Best Local Similarity 74.9%; Pred. No. 9.4e-186; Indels 3; Gaps 1;  
 Matches 790; Conservative 0; Mismatches 262;

Qy 437 AAGAAAGATGGCAACGGACGGCGAGCGGTTGCTGGATATGCCAGGGCTATG 496  
 Db 424 AAGAAATAATGGCAACACAGCCAGAAACCTTGTAGATCCGATTCGGTATG 483

Qy 497 ATCATCTCGATGCCCTGACTCTGATCACTGATCAAGTCAGCTGAGGCTGACGCT 556  
 Db 484 AAATTCTAGATCGCTTAACTTCGATCAAGTTAAGTGAGCTAAAGCTCCATGTAT 543  
 Qy 557 TCCTCGCTGAGCTTATCGTGTGCTGCAAAACAGATCGATCAGCGGTGATCTG 616  
 Db 544 TCTTAGGGTGAATCTATCGTTACTGCTAANGCAATTAAACAGCCTTACATTG 603

Qy 617 GGATCACCGAGCCGGTGGCGCCAGGGGGAGTAATCGCCATTTGGTTAGGTC 676  
 Db 604 GCATTACGAGCAGGTGGCGACCGCTGCGTGTGAGTGGTTAGGAA 663

Qy 677 TGCTGCTCTGAAGGCATGGGACACCGCTGGGTAAGCTGGGGATCCGGCG 736  
 Db 664 TGTTATAGTGAGGGCACTACACTACSGCTCTTTSGGGAGATCCTGTA 723

Qy 737 AAGAGATCAAGTCGTTGATATTTGAAATCCTGGCTATCGTTGGAGGATCA 796  
 Db 724 AGGAATCAAGTCGTTTGTATTTGAAATCTTACGGATTCGGTCRAGGAGATTA 783

Qy 797 ACTTCATCGCTGCCGACGCTGGCGTGGCAGGAATTGTTGTTATCGTGTACG 856  
 Db 784 ACTTTGCTTGCCAACTGCTGCGCAGAATTTGATTAATCGTACAGATAATG 843

Qy 857 CGCTGAGAACGCCCTGGAGATATCATCACTCCGGACGCTTICGATATCGCTGC 916  
 Db 844 CGCTAGAACACGCCCTGAGATATTACACCATGATGTAATCTATPATCGTTG 903

Qy 917 TGGTGAATGGCCCAAGSTGAGCGCTGTTCTACACTCCGGTCAACCGGGCAACAAGA 976  
 Db 904 TAGTGRATGGCTGGCAGGCACTGCTCGATCTGGCTAACGGGGTACAAAN 963

Qy 977 AAAGCGCCTCTGAGATGGCGTGGCGC--AAAGACCTCTGACACAAACGATATA 1033  
 Db 964 AAAGCGGTATTATCTGACGGAGACGCCAAAAGAGCGTTTGATAACGAAGATATAG 1023

Qy 1034 TCGACCAGCTGGAAGCACCAATTGGCGGAAGTC 1068  
 Db 1024 TGAACCAATTAGAAGCAAAATTGGCGGAAGTC 1058

RESULT 6  
 US-09-170-187-7  
 ; Sequence 7, Application US/09170187  
 ; Patent No. 6333745  
 GENERAL INFORMATION:  
 ; APPLICANT: Rather, Philip N.  
 TITLE OF INVENTION: Methods For Screening For Antimicrobials  
 TITLE OF INVENTION: Utilizing aarc And Compositions Thereof  
 NUMBER OF SEQUENCES: 13  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Medlen & Carroll, LLP  
 STREET: 220 Montgomery Street, Suite 2200  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: United States of America  
 ZIP: 94104  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/170, 187  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/827, 190  
 FILING DATE:



QY 203 CGCGCAAGGGTTCAACTCATCAAACASCGGTTAACGGCCTGGCTGCACATCC 262  
 Db 766 CAGCAGAAGCCTTAATTAAATTAACTAACGCGCTGAATGTGCACTGGTNGGGATTC 707

QY 263 ACTTGGACTATCGCATGGCTGAAGTAGGGAAATACGGCTGGATGCTGGTATA 322  
 Db 706 ACTTGGACTATCGCATGGCTGAAGTAGGGAAATACGGCTGGATGCTGGTATA 647

QY 323 ACCCTGGCAATATCGSTAATGAGAGCGTATTCGATGGGGTGACTGCGCCGATA 382  
 Db 646 ACCCAGGTTATATCGGAGCTGAGAGCGTATTCGCCAAGTCGTGATAGTCGTGTCATC 587

QY 383 AAACATTCGATCGTATGGCTTAACGCCGATGCCGAAAGAAGCTGCAGAAA 442  
 Db 586 AACACATTCCTATCGCTATAGGGTCAATGCCGCTACTGGCAAAAGATTCAGAAA 527

QY 443 AGTATGGCACCGACGCCGCGCCGAGGGTGTGTAATCTGCCATGGCTCATGTGATC 502  
 Db 526 AATACGGTGAACCAACACCTGAGCTGAGTCAGTCAAGTCAGTCTAAGGGCGGATCTC 467

QY 503 TCGATCGCTTGACTCTGATCAGTCAGTCAAGTCAGGTGAATCAGCAATGGCAGATGTGATC 562  
 Db 466 TGGACAGGGCTGAATTTCGATCAGTCAGTCAAGTCAGTCTAAGGGCGGATCTC 407

QY 563 CTGTTGAGTCATGTTGCTGGCAAACAGATGATCAGCCTGCACTGSGATCA 622  
 Db 406 CGGTGGCTCTATTCGTTATTGGCCAAATAATGATCACCACITCACTCGTTA 347

QY 623 CCGAAGCCGGTGTGCGCCGAGGGGCACTAAATCCGCACTGGTTAGTCGTC 682  
 Db 346 CAGAAAGGGGGCTCGTTCTGGTTCTAGTAAATCAGCAATTGGCTCTGGT 287

QY 683 TGTCTGAGGCTATCGGCACACGCTGGCGTATCGTGGGGCCATCCGGCTCGAGA 742  
 Db 286 TGGCTGAGGTTATCGGCATAGTTACGTATCTCACTCGGAGATCTCGTGTGAGGA 227

QY 743 TCAAAGTCGGTTCTGATATTTGAAATTCGCTGGTACCGATCCGTTGGAGGATCACTCA 802  
 Db 226 TGAAGTCGGTTCTGATATTCTAAATCCTGTTACGGATCCGCTTACGGATCACTTA 167

QY 803 TCGCCPCCGACCTCTTCAGCCAAAGATTCGCTTACGGTTACGGCTGG 862  
 Db 166 TTGCTGPGCCAACTCTTCAGCCAAAGATTCGCTTACGGTTACGGCTGG 107

QY 863 ACCAAAGCCCTGGAGATATCTCACTCCGATGGCTTCCATATCGGCTGGTGA 922  
 Db 106 AGCAGCGCTCGAGATATTATCAGCCGATGAGTCCTPATATGGTGTGAGTA 47

QY 923 ATGGCCAGGGCTGGTACGGGCTGTTTACACTGGCTCACGGCG 967  
 Db 46 ATGGCCGGGTTAACGCCGAGGTTCIACTTGGTGTGGCTGGCG 2

**RESULT<sup>8</sup>**  
 US-09-170-187-10c  
 Sequence 10, Application US/09170187  
 Patent No. 6,383745  
 GENERAL INFORMATION:  
 APPLICANT: Rather, Philip N.  
 TITLE OF INVENTION: Methods For Screening For Antimicrobials  
 NUMBER OF SEQUENCES: 13  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Medgen & Carroll, LLP  
 STREET: 220 Montgomery Street, Suite 2200  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: United States of America  
 ZIP: 94104  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/170,187  
 FILING DATE:  
 CLASSIFICATION:  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: 08/827,190  
 FILING DATE: 08/27/1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Carroll, Peter G.  
 REGISTRATION NUMBER: 32,837  
 REFERENCE/DOCKET NUMBER: CASE-02443  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 705-8410  
 FAX: (415) 397-8338  
 INFORMATION FOR SEQ ID NO: 10:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 886 base pairs  
 STRANDEDNESS: double  
 TOPOLOGY: 886 base pairs  
 MOLECULE TYPE: DNA (genomic)  
 US-09-170-187-10

Query Match 47.8%; Score 534.6; DB 4; Length 886;  
 Best Local Similarity 75.3%; Pred. No. 6.7e-159; Mismatches 219; Indels 0; Gaps 0; Matches 666; Conservative 0; Mismatches 219; Indels 0; Gaps 0;

QY 83 TCGCCGCTACGTCATGACCAATACGGTACAGACGACGTCGAGCAACGGTCATCAAA 142  
 Db 886 TTGCTGTCACAATCTATGACGATAACCGCAGACGGATGTTGAAAGCCACATGCGCAAA 827  
 QY 143 TCAAGGGCTGGAACCGGGTGGCGCTGATATCGCTGGTATACCGTACCGACGAGCG 202  
 Db 826 TCCATCACTTGAGGTGAGTGTGATATGGCTCGTGTGTCCTCPACGATGATG 767

QY 203 CGGCACAAAGCTCAACATCTCAACAGCAGGTACGGCGCTGATGTCCTCGGATCC 262  
 Db 766 CAGCAGAAGCCTTAATTAAATTACGCAAGCCGTTGATGTCCTCGGATTC 707

QY 263 ACTTCGACTATCGCATGGCTATGAGAGCGTATTCGCTGGTGAATATGGTGTGACTGCTGATT 322  
 Db 706 ACTTGGACTACCTGATCTGGCTGAAGTACGGCTCATGGTGTGACTGCTGCGCGATA 647

QY 323 ACCCTGGCAATATCGCTATGGCTATGAGAGCGTATTCGCTGGTGAATATGGTGTGACTGCTGCGCGATA 382  
 Db 646 ACCCAGGTTATATCGGCAAGACGCTTACGGCTCACGGCTAACAGGATATCCGATTC 587

QY 383 AAACATTCGCTTACGGCTATGGCTAACGGCTGATGGCTCATGGGGCTACGGAAAGATCTGAGAAA 442  
 Db 586 ACACATTCCTATCCGTTAGGGCTCATGGGGCTACGGAAAGATATCCGAGAA 527

QY 443 AGTATGGCACCGACGCCGAGGGCTGGTGTGAGCTGCAATCTGCATGGCTCATGTC 502  
 Db 526 AATACGGTGAACACCTGAGCTGAGCTGATGGTGAATCTGCACATGGCACATGTGATC 467

QY 503 TCGATGGCTGACTCTGATGCACTGAGCTCAAGCAGGGTGAAGGCGCTGACGCTCTCG 562  
 Db 466 TGGACAGGGCTATGGCTTATGGCCAAATAATGATCAGCTAACGCGTGGATCTCTG 407

QY 563 CTGTTGAGCTATCGCTTCTGGCTTATGGCCAAATAATGATCAGCTAACGCGTGGATCTCTG 622  
 Db 406 CGGTGGCTCTATCGCTTATGGCCAAATAATGATCAGCTAACGCGTGGATCTCTG 347

QY 623 CGGAAGGGGGTGGCTGGCGAGCGGGCGAATGGCTAACAGCTAACGGCTGGATCTGGTATA 682  
 Db 346 CAGAAGGGGGTGGCTGGCTGGCGAGCGGGCGAATGGCTAACAGCTAACGGCTGGATCTGGT 287

QY 683 TGTCTGAGGCTATGGCTGGCGAGCGGGCGAATGGCTAACAGCTAACGGCTGGATCTGGTATA 742  
 Db 286 TGGCTGAGGTTATGGCTGGCGAGCTGGCTAACAGCTAACGGCTGGATCTGGTATA 227

QY 743 TCAAAGTCGGTTCTGATATTGAAATCCGCTATCCGTTCCGGAGGCTAACATCA 802

RESULT 9  
US 08-827-190-9/C  
Sequence 9, Application US/08827190  
; Patent No. 585367

GENERAL INFORMATION:  
 APPLICANT: Rather, Philip N.  
 TITLE OF INVENTION: Methods For Screening For Antimicrobials  
 NUMBER OF SEQUENCES: 13  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Medlen & Carroll, LLP  
 STREET: 220 Montgomery Street, Suite 2200  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: United States of America  
 ZIP: 94104

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS DOS

SOFTWARE: PatentIn Release #1.0, Version #1.3.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/827,190  
 FILING DATE:  
 CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
 NAME: Carroll, Peter G.  
 REFERENCE DOCKET NUMBER: CASE-02443  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 397-8338  
 TELEFAX: (415) 705-8410

INFORMATION FOR SEQ ID NO: 9:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 886 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: unknown  
 MOLECULE TYPE: DNA (genomic)

US-08-827-190-9

Query Match 24.0%; Score 269; DB 2; Length 886;  
 Best Local Similarity 56.5%; Pred. No. 4; 8e-75; ;  
 Matches 500; Conservative 0; Mismatches 385; Indels 0; Caps 0;

QY 83 TCGCCGCTACAGTCATGCCAATACGGTAGCAGACAGCTGCAATCAA 142  
 QY 886 TTGTCTATCCAAGCATGACAACAAACATGATGTTAGAGAACGGTTGGGAA 827  
 QY 143 TCAAGGGCTGGACCGCTGGCGCTGATATCGTCCGGTATCGTACCGAGATGGACG 202  
 Db 826 TTACCGTTGGCTGAGGCCGATGGCAATTCGCTGGGATGAACGGG 767  
 Qy 203 CGGAGAAGCTTCAACTCATCACAGCAGGTTACGTCGCCGCTGCTGACATCC 262  
 Db 766 CGGAAACGCCATTGGGATATAAAAGCCATTTCATTCCTCGTGTGACATAC 707

RESULT 9  
US 09-170-187-9/C  
Sequence 9, Application US/09170187  
; Patent No. 638745

GENERAL INFORMATION:  
 APPLICANT: Rather, Philip N.  
 TITLE OF INVENTION: Methods For Screening For Antimicrobials  
 NUMBER OF SEQUENCES: 13  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Medlen & Carroll, LLP  
 STREET: 220 Montgomery Street, Suite 2200  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: United States of America  
 ZIP: 94104

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.3.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/170,187

US-09-170-187-9

Query Match 26.3%; Score 322; DB 3; Length 886;  
 Best Local Similarity 56.5%; Pred. No. 4; 8e-75; ;  
 Matches 500; Conservative 0; Mismatches 385; Indels 0; Caps 0;

QY 263 ACTCGACTATCCGATTCGGCTGAAAGPAGCGGAATCGGGCTCGATGTTCTCGCTTATA 322  
 Db 706 ATTCGATTAATCTTGCTGAAAGCCATGAGCGGCAGATRAATCCGAATCA 647  
 QY 323 ACCCTGGATAATGGTTATGAAGAGCGTATTGCGATGGGGTACTGTGGGGTACTGTGGGGCATA 382  
 Db 646 ACCCGGAAATATGGCGGGCGAAAGTGAAGGGTTAAAGGCCAAGAC 587  
 QY 383 AAACATTCGATCGTATTGGGTTAACGCCGATGCTGGAAAAGATCTGCAAGAAA 442  
 QY 586 AAGGCATTCGATCAGAACCTGGGTAACAGCCGTTATGGAAAAGGATTTAGAAA 527  
 QY 443 ACTATGGGAAACGACGCCGCGAGGGTGTGGAATGCCATGCGTCATGTGATCATC 502  
 Db 526 AAATACGGTTATCGACTGCCGATGAAGCAGCTCATCACATTAATTC 467

QY 503 TCGATGCCGAACTCGATCAGTCAGTCAAAGTAGCGTAAAGCTGAGGCTCTGACGGCTCTC 562  
 Db 466 TTGAGGATCTGATTCACGTTATTAGTGTAGCTGAAGGCTCTGACGGTAACCTT 407  
 QY 563 CTGTTGACGTTCTATCGTTGCTGCAACACAGATCGATCGATGCCATGGTTAGCTG 622  
 Db 346 CCGAGTCAGGAACACTGTTGCGCAGACTAACAGGCGCAGACTCGGGCCATT 287  
 QY 683 TGCTGAGCATGGCACACGCTGCOCGATTCGCTCGGCCGATCCGGTCAAGAGA 742  
 Db 286 TAAGCAAAGGCATGGAACACATGCGCATTCAGCTAACGCCAGACCCCTGAGAAGGG 227  
 QY 743 TCAAGTGGTTGATATTGTAATCGTCTCGTATCCGTTGCGAGGGATCACTCA 802  
 Db 226 TAAAGTAGCAAGGGACTCTGCAATCTTGGCTTAGGCCATGCTCCACGCTCA 167

QY 803 TCGCCGCTCCGACCTGTTGCGCTCAGSARATTTGATGTTACGGTAGGTTAACGCCCTG 862  
 -Db 166 TCTCATGCCGACTTGGGGCGTATTGAGATGATGATCTAACGATGCGCATGAACTGG 107

QY 863 ACCAACCGCTGGAGATATCATCAGTCGATGACGCTTGCATTATGGCTAGGGCTG 922  
 Db 106 AAGATATTCTAAGATAAAAGGCCGATTAAGTGTGCTCTGGCTGGCTG 47

QY 923 ATGGCCAGGTGGCCGCTGGTTCTACCTGGCTCACCCGGG 967  
 Db 46 AGGGACCTGGAAAGCGAGAGCTGATTCGGAAATCGGGGG 2





Db 1312 TCGGCAGCCACCTGCCGATGGGGCCCGCAGCTGGGTGCTGCTGAGTAGGAGTCG 1253  
Qy 639 GCGCAGGGGGCACTTAAATCGCCATTGGTTAGGCTCTGCTGCTGAGGATCGG 698  
Db 1252 GCGTGGCCAGCACTCACCCTGGCCATGCCCTGCGATGAGGGATCGC 1193  
Qy 699 CGACAGCTGGCGTATCGCTGGCGAGAAGATAAAGTGGGTTCGA 758  
Db 1192 CGGTGGCCAGCGCGATGTCGTCGAGCGCGTAGACTCGTCA 1133  
Qy 759 TAT 761  
Db 1132 TCT 1130

RESULT 14  
US-09-221-017B-1017/c  
; Sequence 1017, Application US/09221017B  
PATENT NO. 644799  
GENERAL INFORMATION:  
APPLICANT: ROSS, Bruce C.  
TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF  
NUMBER OF SEQUENCES: 1120  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 755 PAGE MILL ROAD  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304-1018  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: FASTSEEK for Windows Version 2.0b

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/221,017B  
FILING DATE: 23-DEC-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PP1182  
FILING DATE: 31-DEC-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PP1546  
FILING DATE: 30-JAN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PP2911  
FILING DATE: 09-APR-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/AU98/01023  
FILING DATE: 10-DEC-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: MONROY, Gladys H  
REGISTRATION NUMBER: 32,430  
REFERENCE/DOCKET NUMBER: 27340-20021.00  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-813-5600  
TELEFAX: 650-494-0792  
TELEX: 706141  
INFORMATION FOR SEQ ID NO: 1017:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1665 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: UNKNOWN  
ORIGINAL SOURCE: ORGANISM: PORPHYROMONAS GINGIVALIS  
FEATURE: NAME/KEY: misc feature  
LOCATION: 1..1665

RESULT 15  
US-09-103-840A-1/c  
; Sequence 1, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; FILE REFERENCE: 24366-2007.00  
; CURRENT APPLICATION NUMBER: US/09/103,840A  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 4411529  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; OTHER INFORMATION: H37RV  
; US-09-103-840A-1

Query Match 34%; Score 38.2; DB 4; Length 4411529;  
Best Local Similarity 53.5%; Pred. No. 3.7; Mismatches 462; Indels 3; Gaps 1;  
Matches 534; Conservative 0; Mismatches 462; Indels 3; Gaps 1;

Qy 4 9 GTGGGAATGCGGATTGCGATGGTGTGCTGCCATGCGCTACAGTCCATGACCAATAGC 108  
Db 3180453 GTCGGCAACGTCGGTGGCGAGTGGACATCCGGTCTGGTCAATCGATGCAACCC 3180394  
Qy 10 9 CGTAGACAGACGCTGAAACGCGTCAATCAAGCGCTGAGGCGCGAGGGTCAACTCATCAA 228  
Db 3180393 AAACCCACGACGTCACATGACATGGACAAATGCCGAGCGACCGCGATGCC 3180334  
Qy 16 9 GATACGTCGTGATCCGTACCGCGATGCGAGAAGGGTCAACTCATCAA 228  
Db 3180333 GACATCGTGGGGTGCGCCTGGCGCCAGGAGACGCGCGACGCGCTGCGAGATGCC 3180274  
Qy 22 9 CAGCAGTTCACGTCGGTGGTGTGCTGACATCCACTTGACTATCGCATGGCGCTGAAA 288  
Db 3180273 CGCGCACAGACGATGCCGTGACGCGACATACATTCCAGCGCGTACATATTCCC 3180214  
Qy 28 9 GTAGCGATTAACGGGTCAATGTCGCTGATTAACCCGGCGACATCGCGCTGCGAGATGCC 3180214  
Db 3180213 GCCATCGACGCTGGATGTCGCCGCGTGGGGTCAACCGGGGACATCGAGGTTGAC 3180154  
Qy 34 6 GAGGTATGCGATGGTGGTGAATGTTGCGGCGCATAAACATTCGATCCGATTGCG 405  
Db 3180153 GGCGGGTGGTGGAGGTCGCCAACGGCGCGGGTGGCGGATCCCCATCCGAATCGGT 3180094

Mon May 12 08:44:41 2003

us-09-921-992-3.rni

Page 12

QY	406	GTTAACCCGGGATCGCTGGAAAAGATCTGCAAGAAAGTATGGCGACCGAGCCACGCCAG	465
Db	3180093	GTCAAACGCCCGATTGCTGTCGAGAACCGTCATGGAGAATGTCGCAAGGCCACGCCAG	3180034
QY	466	GCGTGTCTGGAATCTCCCATCGTCAGTGTGATCAGTCGATCGCCTGACTCGATGAG	525
Db	3180033	GCGCTGTCTGAGTCGCGCTGTTGGGGAGCTTGCTTTCGAGGAGATGGCTTCGGTAC	3179974
QY	526	TTCAGACTCAGCTGAAGGGCTGACGCTTCCTCGCTGAGCTTATCGTTGCTG	585
Db	3179973	ATCAAGATCAGCTCAGGAAACGACCCGGTGTGATGGTCGCCCTACAGCTGCTT	3179914
QY	586	GCAAACAGATGATGATCAGCGCTGTCATCTGGGATCACCGAAGCCGGTGTGGCGAGC	645
Db	3179913	GCTGCACGGTGTGACTACGGACTGCACCTCGGTGTCACCGAGGCCGCTTCAG	3179854
QY	646	GGGGCACTAAATCGCCAATGTTAGTCGCTGCTGCTGTTGTCGAAGGATCACGGAC	705
Db	3179853	GCCACCATCAAGTCCCGCTGGCTTCGGCTGGCGCTGTTGTCGGGGCAATAGGCAC	3179794
QY	706	CTGGCGCTATCGCTGGGGGGATCGCTGAAGAGATCAAATCGTTCTGATATTG	765
Db	3179793	ATCCGGGTGTTGTTGRCGGGCCCGGGCTGAGGAAGTCAGGTGGCAATCAGGTTTC	3179734
QY	766	AATCGCTCGGATACCGTTGGAGGGATCAACTCATCGCTGCCGACCTTCGGGT	825
Db	3179733	GAGTCGTTGAACCTGGGGGGCGCTTCGAGATCGTCCTGCGCCCTGGCTG	3179674
QY	826	CAGGAATTGATGTTATCGTCAGCTTACCGCTGGAGAACGCCGCTGAAATATATIC	885
Db	3179673	GCGCAAGTCGAGCTTACCGCTGGAGAACCGGGCTGAGTAACCGCGGCCCTGATGATGTCGAT	3179614
QY	886	ACTCGATGAGCTTCGATTTCTGGCTCGCTGGTGAATGGCCAGGGAGGGCTGTT	945
Db	3179613	GTGCGCTGGTGCAGGGTGTGATGGGGTGTGTCATGGHOCGGGTGAACGACCTGAG	3179554
QY	946	TCTACACTCGGCTCACGGGGGGCAAAAGAAAAGCGCCCTATGAGATGGTGGC	1005
Db	3179553	GCGGACCTGGGGTGGCTGGCTGGCGACAGTCAGTCAGATGTCGAGACGGT	3179494
QY	1006	AAAGACGGCTTGGACAACAAAGATGATGTCGAGCTG 3179495	1044
Db	3179493	ATCAAGACCGTCCCCGAGGACAGATGTCGAGACGGT	3179455

Gencore version 5.1.4\_p5\_4578  
 Copyright (c) 1993 - 2003 Compugen Ltd.

## OM nucleic - nucleic search, using sw model

Run on:

May 10, 2003, 16:41:08 ; Search time: 197 seconds

(without alignments)

7058.816 Million cell updates/sec

Title: US-09-921-992-3  
 Perfect score: 1119  
 Sequence: 1 atgcataacccaggctccaaat.....ttcaagcagggtgaaaaataa 1119

Scoring table: IDENTITY\_NUC  
 Gapop 10.0 , Gapext 1.0

Searched: 783854 seqs, 621352466 residues

Total number of hits satisfying chosen parameters: 1567708

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_NA,\*

1: /con2\_6/ptdodata/2/pubpna/us07\_pubcomb\_seq.\*  
 2: /cgn2\_6/ptdodata/2/pubpna/pcr\_new\_pub\_seq.\*  
 3: /con2\_6/ptdodata/2/pubpna/us06\_pubcomb\_seq.\*  
 4: /cgn2\_6/ptdodata/2/pubpna/us07\_new\_pub\_seq.\*  
 5: /con2\_6/ptdodata/2/pubpna/pcrus\_pubcomb\_seq.\*  
 6: /con2\_6/ptdodata/2/pubpna/us08\_new\_pub\_seq.\*  
 7: /cgn2\_6/ptdodata/2/pubpna/us08\_pubcomb\_seq.\*  
 8: /cgn2\_6/ptdodata/2/pubpna/us09\_new\_pub\_seq.\*  
 9: /cgn2\_6/ptdodata/2/pubpna/us09\_pubcomb\_seq.\*  
 10: /cgn2\_6/ptdodata/2/pubpna/us10\_new\_pub\_seq.\*  
 11: /cgn2\_6/ptdodata/2/pubpna/us10\_pubcomb\_seq.\*  
 12: /cgn2\_6/ptdodata/2/pubpna/us10\_pubcomb\_seq.\*  
 13: /cgn2\_6/ptdodata/2/pubpna/us60\_new\_pub\_seq.\*  
 14: /cgn2\_6/ptdodata/2/pubpna/us60\_pubcomb\_seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	1119	100.0	1119	US-09-921-992-3
2	403.6	36.1	640681	Sequence 3, Appli
3	293.4	26.2	1083	Sequence 1, Appli
4	243.8	21.8	1164	Sequence 1692, Appli
5	239.4	21.4	1134	Sequence 110, Appli
6	211.8	18.9	1180	Sequence 2205, Appli
7	70.6	6.3	584	Sequence 215, Appli
8	68.4	6.1	596	Sequence 21, Appli
9	68.4	6.1	670	Sequence 23, Appli
10	56.8	5.1	348	Sequence 22, Appli
11	52.6	4.7	293	Sequence 6102, Appli
12	52.6	4.7	456	Sequence 25, Appli
13	50.5	4.5	403	Sequence 195, Appli
14	48.8	4.4	564	Sequence 21, Appli
15	48.4	4.3	379	Sequence 23, Appli
16	48.4	4.3	528	Sequence 16, Appli
17	48.4	4.3	250	Sequence 15, Appli
18	48.2	4.3	353	Sequence 1, Appli
19	4.3	353	10	Sequence 39, Appli
20				Sequence 33, Appli

## ALIGNMENTS

## RESULT 1

US-09-921-992-3

; Sequence 3, Application US/09921992

; Patent No. US20030069426A1

## ; GENERAL INFORMATION:

; APPLICANT: Boronat, Albert;

; APPLICANT: Campos, Narciso;

; APPLICANT: Rodriguez-Concepcion, Manuel;

; APPLICANT: Rohmer, Michael;

; APPLICANT: Seeman, Myriam;

; APPLICANT: Valentim, Henry E.;

; APPLICANT: Venkataswamy, Venkatesh, Mylavapu

; TITLE OF INVENTION: Methy-D-Erythritol Phosphate Pathway Genes

; FILE REFERENCE: 16516.107/35-21(51897)US

; CURRENT APPLICATION NUMBER: US/09/921,992

; CURRENT FILING DATE: 2001-08-06

; PRIORITY APPLICATION NUMBER: US 60/223,483

; PRIOR FILING DATE: 2000-08-07

; NUMBER OF SEQ ID NOS: 85

; SEQ ID NO 3

; LENGTH: 1119

; TYPE: DNA

; ORGANISM: Escherichia coli

; FEATURE: Best Local Similarity

; NAME/KEY: CDS

; LOCATION: (1)..(1119)

; US-09-921-992-3

Query Match Score 100.0%; DB 10; Length 1119;

Best Local Similarity 100.0%; Pred. 0%; Mismatches 0;

Matches 1119; Conservative 0; Indels 0; Gaps 0;

; Sequence 1, Appli

; Sequence 2, Appli

; Sequence 3, Appli

; Sequence 4, Appli

; Sequence 5, Appli

; Sequence 6, Appli

; Sequence 7, Appli

; Sequence 8, Appli

; Sequence 9, Appli

; Sequence 10, Appli

; Sequence 11, Appli

; Sequence 12, Appli

; Sequence 13, Appli

; Sequence 14, Appli

; Sequence 15, Appli

; Sequence 16, Appli

; Sequence 17, Appli

; Sequence 18, Appli

; Sequence 19, Appli

; Sequence 20, Appli

; Sequence 21, Appli

; Sequence 22, Appli

; Sequence 23, Appli

; Sequence 24, Appli

; Sequence 25, Appli

; Sequence 26, Appli

; Sequence 27, Appli

; Sequence 28, Appli

; Sequence 29, Appli

; Sequence 30, Appli

; Sequence 31, Appli

; Sequence 32, Appli

; Sequence 33, Appli

; Sequence 34, Appli

; Sequence 35, Appli

; Sequence 36, Appli

; Sequence 37, Appli

; Sequence 38, Appli

; Sequence 39, Appli

; Sequence 40, Appli

; Sequence 41, Appli

; Sequence 42, Appli

; Sequence 43, Appli

; Sequence 44, Appli

; Sequence 45, Appli

; Sequence 46, Appli

; Sequence 47, Appli

; Sequence 48, Appli

; Sequence 49, Appli

; Sequence 50, Appli

; Sequence 51, Appli

; Sequence 52, Appli

; Sequence 53, Appli

; Sequence 54, Appli

; Sequence 55, Appli

; Sequence 56, Appli

; Sequence 57, Appli

; Sequence 58, Appli

; Sequence 59, Appli

; Sequence 60, Appli

; Sequence 61, Appli

; Sequence 62, Appli

; Sequence 63, Appli

; Sequence 64, Appli

; Sequence 65, Appli

; Sequence 66, Appli

; Sequence 67, Appli

; Sequence 68, Appli

; Sequence 69, Appli

; Sequence 70, Appli

; Sequence 71, Appli

; Sequence 72, Appli

; Sequence 73, Appli

; Sequence 74, Appli

; Sequence 75, Appli

; Sequence 76, Appli

; Sequence 77, Appli

; Sequence 78, Appli

; Sequence 79, Appli

; Sequence 80, Appli

; Sequence 81, Appli

; Sequence 82, Appli

; Sequence 83, Appli

; Sequence 84, Appli

; Sequence 85, Appli

; Sequence 86, Appli

; Sequence 87, Appli

; Sequence 88, Appli

; Sequence 89, Appli

; Sequence 90, Appli

; Sequence 91, Appli

; Sequence 92, Appli

; Sequence 93, Appli

; Sequence 94, Appli

; Sequence 95, Appli

; Sequence 96, Appli

; Sequence 97, Appli

; Sequence 98, Appli

; Sequence 99, Appli

; Sequence 100, Appli

121 GTCGAGACGACGGTCATCAATCAGGCGTTGGAACGGCTTGGCGTGTATCGTCGT 180  
 181 GTATCGTACCGACGATGGGGAGAAGCGTCAAACTCATCAAACACAGGTTAAC 240  
 181 GTATCGTACCGACGATGGGGAGAAGCGTCAAACTCATCAAACACAGGTTAAC 240  
 241 GTGCCGCTGGGCTACATCACTCGACTATGGCATGGCTGAAAGTACGGTAC 300  
 241 GTGCCGCTGGGCTACATCACTCGACTATGGCATGGCTGAAAGTACGGTAC 300  
 301 GGCGRGATATTCTGGTATTAACCTGGCATATGGTAAAGAAGCGTATTCGATG 360  
 301 GGCGRGATATTCTGGTATTAACCTGGCATATGGTAAAGAAGCGTATTCGATG 360  
 361 GTGGTGACTCTGCGCGATAAACATTCGATCGTATGGITAAGCCGGATCG 420  
 361 GTGGTGACTCTGCGCGATAAACATTCGATCGTATGGITAAGCCGGATCG 420  
 421 CTGGAAAAGACTCGGAGAAAGTATGGCCACCGAGGCCAGGGTCTGGATCT 480  
 421 CTGGAAAAGACTCGGAGAAAGTATGGCCACCGAGGCCAGGGTCTGGATCT 480  
 481 GCCATCGCTCACTGTTGATCACTCGATGCCAACCGAGGCCAGGGTCTGGATCT 540  
 481 GCCATCGCTCACTGTTGATCACTCGATGCCAACCGAGGCCAGGGTCTGGATCT 540  
 541 AAAGGCTCTGCTTCCTCTGCTGTTGAGCTTATGGTGTCTGGCAAAGCTGAT 600  
 541 AAAGGCTCTGCTTCCTCTGCTGTTGAGCTTATGGTGTCTGGCAAAGCTGAT 600  
 601 CAGCCCTTGTCTGGGATACCGAGCCGGCTGGCCSAGGGGCTTAATCC 660  
 601 CAGCCCTTGTCTGGGATACCGAGCCGGCTGGCCSAGGGGCTTAATCC 660  
 661 GGCATGGTTAGCTCTGCTGTCGAAGACGATGGCGTATGGCTG 720  
 661 GGCATGGTTAGCTCTGCTGTCGAAGACGATGGCGTATGGCTG 720  
 721 GCGGGGATCGGTCAAAGAACATCAAGTCGCTTGTGATTTGAAATCGGTATC 780  
 721 GCGGGGATCGGTCAAAGAACATCAAGTCGCTTGTGATTTGAAATCGGTATC 780  
 781 CGTCTGGAGGAGTCACTCTATGGCTGGCAGGTTGGCTCAGGATTGGCT 840  
 781 CGTCTGGAGGAGTCACTCTATGGCTGGCAGGTTGGCTCAGGATTGGCT 840  
 841 ATCGGTACGGTTAACCGCTGAGCAACGCCGCTGGAGATATCATCACTCCGAGCTGTT 900  
 841 ATCGGTACGGTTAACCGCTGAGCAACGCCGCTGGAGATATCATCACTCCGAGCTGTT 900  
 901 TCGATGATCGCTGGCTGGAATGGCCCAAGTGGCTTACCGCTGGCTGCT 960  
 901 TCGATGATCGCTGGCTGGAATGGCCCAAGTGGCTTACCGCTGGCTGCT 960  
 961 ACCGGGGCACAGAAAGGGCCCTATAGATGGCTGGCAAGACCGCTGGAC 1020  
 961 ACCGGGGCACAGAAAGGGCCCTATAGATGGCTGGCAAGACCGCTGGAC 1020  
 961 ACCGGGGCACAGAAAGGGCCCTATAGATGGCTGGCAAGACCGCTGGAC 1020  
 1021 ACAACGATATGATGACCACTGGAGCAAGCCGCTATGGCTGGCAAGACCGTCTGGAC 1080  
 1021 ACAACGATATGATGACCACTGGAGCAAGCCGCTATGGCTGGCAAGACCGTCTGGAC 1080  
 1081 GAAGGGCTGCAATGGCTAGCGAGTTAAATA 1119  
 1081 GAAGGGCTGCAATGGCTAGCGAGTTAAATA 1119  
 1081 GAAGGGCTGCAATGGCTAGCGAGTTAAATA 1119

RESULT

US-09-790-988-1

1

; Sequence 1, Application US/09790988  
; Patent No. US2002012768A1  
; GENERAL INFORMATION:

Qy 799 TTCACTGGCTGCCGACCTGTTGGCTCAGGAAATTGATGTTACGGTAAACGGC 858  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Db 315073 TTATGTGCTGTCCTACTGTTCTAGACAGAATTGATGTAATACGATAATCAA 315132  
 Qy 859 CTGGAGCAACGCCCTGGAAGATATCAGTCACTCGACTCCATAGATTTGCGCTG 918  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Db 315133 CTTAGAAAAAATCTGAGATATCTCGACTCCATAGATTTGCGCTG 315192  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Qy 919 GTGAATGGCCAGGAGGGCTGTTACACTGGGGTCAAGGGCACACAGAA 978  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Db 315193 GTTAAATGGAAATGGGATCTAATAGCAACTTGGGTCTAGGAAACTCATAGAA 315252  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Qy 979 AGCGGCCTCTATGANGATGGCT--GGCAAGACGACCCGCTGGACACACGATATGATC 1035  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Db 315253 AGTGCATTTRATGAAGACGCGAGTAAGACAAAAAGAAAAATAAAACGAAATTATA 315312  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Qy 1036 GACCGAGCTGAAAGCACGGCATTCGTCGAA 1065  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Db 315313 GAAAATGGAAATTAAATTCGAAAAA 315342  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 RESULT 3  
 US-09-974-300-1692  
 ; Sequence 1692, Application US/09974300  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Berka, Randy M.  
 ; TITLE OF INVENTION: Methods For Monitoring Multiple Gene  
 ; TITLE OF INVENTION: Expression  
 ; FILE REFERENCE: 10085\_500-US  
 ; CURRENT FILING NUMBER: US/09/974, 300  
 ; CURRENT FILING DATE: 2001-10-05  
 ; PRIOR APPLICATION NUMBER: 09/680, 598  
 ; PRIOR FILING DATE: 2000-10-06  
 ; PRIOR APPLICATION NUMBER: 60/279, 526  
 ; PRIOR FILING DATE: 2001-03-27  
 ; NUMBER OF SEQ ID NOS: 8481  
 ; SOFTWARE: fastSEQ for Windows Version 4.0  
 ; SEQ ID NO: 1692  
 ; LENGTH: 1083  
 ; TYPE: DNA  
 ; ORGANISM: Bacillus licheniformis  
 ; US-09-974-300-1692  
 ; Query Match 26.2%; score 293.4; DB 10; Length 1083;  
 ; Best Local Similarity 56.6%; Pred. No. 1.2e-90; Mismatches 543; Conservative 0; Indels 0; Gaps 0;  
 ; Matches 416; Gaps 0;  
 Qy 22 CACCGTAGAACATCACACGCTTACGGTGGAAATGGCGATGGCGATGGCTCCC 81  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Db 21 CATGGCTAACAAACGGCTCCGGTTAACGGACCTTAACTAACAGCGCAATACGAA 80  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Qy 82 ATCGCGCTACGGCATGACCATACGGCTACGACGACCTCGAAGCAACGGTCAATCAA 141  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Db 81 GTCTGCATTCGAAGCATGACACAAACACATGACCTGGCTGTTAACGGCGAA 140  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Qy 142 ATCAAGGCGCTGGAAACGGCTGGCGCTGATATCGTCCGGTATCGTGTACCGAC 201  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Db 141 ATCAACAGACTCGCGGAAGCAGATGTCATCGTCCGGCTGGCTGTGACGG 200  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Qy 202 GGCGCAGAACGGCTGAAACTCATCAACACGGGTTAACCTGGCCCTGGCGACATC 261  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Db 201 GCTGCCGACGCCATCCAGATCAACAGGATATCCATCCCTTGTCGTGATTT 260  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Qy 262 CACTGACACTCGCATGGCTGAAAGTAGGGATACGGCTGGATCTGGTATT 321  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Db 261 CATTTCACATATAATTGGCATATAAGGGATCGAAGGGGAGGGCGATATAATCGCATC 320  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Qy 322 AACCCCTGGCAATATGGGTAATGAGAAGCGCTTATCGATGGTGTGAGCTGCGCGAT 381  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
 Db 321 ATCCCGGTTACATGCCGCCGAGAAGGTGAGCGGCTGTCACCGCAGCGAGGA 380  
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 4  
 US-09-712-363-110  
 ; Sequence 110, Application US/09712363  
 ; Patent No. US/0020164588A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Eisenberg, David  
 ; APPLICANT: Rottstein, Sergio H.  
 ; APPLICANT: Marcotte, Edward M.  
 ; TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND  
 ; FILE REFERENCE: 07419-032001  
 ; CURRENT APPLICATION NUMBER: US/09/712, 363  
 ; CURRENT FILING DATE: 2000-11-13  
 ; PRIOR APPLICATION NUMBER: PCT/US00/02246  
 ; PRIOR FILING DATE: 2000-01-28  
 ; PRIOR APPLICATION NUMBER: 60/179, 531  
 ; PRIOR FILING DATE: 2000-02-01  
 ; PRIOR APPLICATION NUMBER: 60/117, 844  
 ; PRIOR FILING DATE: 1999-01-29  
 ; PRIOR APPLICATION NUMBER: 60/118, 206,  
 ; PRIOR FILING DATE: 1999-02-01  
 ; PRIOR APPLICATION NUMBER: 60/126, 593  
 ; PRIOR FILING DATE: 1999-03-26  
 ; PRIOR APPLICATION NUMBER: 60/134, 093  
 ; PRIOR FILING DATE: 1999-05-14  
 ; PRIOR APPLICATION NUMBER: 60/134, 092  
 ; PRIOR FILING DATE: 1999-05-14  
 ; PRIOR APPLICATION NUMBER: 60/165, 124  
 ; PRIOR FILING DATE: 1999-11-12  
 ; PRIOR APPLICATION NUMBER: 60/165, 086  
 ; PRIOR FILING DATE: 1999-11-12

NUMBER OF SEQ ID NOS: 292  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 110  
; LENGTH: 1164  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; US-09-712-363-110

Query Match 21.8%; Score 243.8; DB 9; Length 1164;  
Best Local Similarity 53.5%; Pred. No. 1.7e-73; Mismatches 0;保守性 1;  
Matches 534; Conservatve 0; Mismatches 462; Indels 3; Gaps 1;  
Db 49 GTTGGGAATGTGCGCATGGCCTGGTCTCCATCCCGTGAGTCATGCCAATAGC 108  
Db 76 GTCGCACAGTCGGCTGGCACTGGACATCGACATCGATGTCGACACC 135  
Qy 109 CGPACGACAGAGTCGAGCAAGGTCATCAATCAAGGCTGGAACGGTGGCT 168  
Db 136 AAAACCCCACGACTGCAACTCGACATGCAACAAATCGCCGACTGACCGCG 195  
Qy 169 GATATCGTCGCTGATCGTACCGACATGGAGCAGANGCGTCAAACCTCAA 228  
Db 196 GACATCTGGGGGGCTGCCGCCAGGAGGACGCCGACGCCCTGGCGAGATCCC 255  
Qy 229 CAGCAGGTTAACGTGCGCTGTTGGCTGACATCACCTGACTATCGATGGCTGAA 288  
Db 256 CGGCACAGCCAGATCCCGCTGGTAGLTCGGACATACATTCCAGCGCTACATTGCC 315  
Qy 289 CTAGCGGAAATAGGGCGTGTGATGTCGCGTATTAACCTGCGAATTC 345  
Db 316 GCATCGACGCTGGATGTCGCGGTGGGTCACCGGGCACATCAAGGTTGAC 375  
Qy 346 GACCGTATTCGCTGGTGTGACTGTCGCGGATAAAACATCGATCGTATGGC 405  
Db 376 GCGCGGGGGGGAGGGGGCGCCAGGGCGCGGGTGGCGCGGATCCCGATCCG 435  
Qy 406 GTCACCGCGGATGCGCAGAAAAGATCTGCGAAGAAAGTAGTGGGACCGGCCAG 465  
Db 436 GTCACACGCCGGTTCGCTGACACACGGTTCATGGAGATGTCGACAMGCCAGGCCAG 495  
Qy 466 GCGTTGGTGGAACTCTGCGCATGGCTGTCATCTGCGACTCTGCGATCAG 525  
Db 496 GCGCTGGTTGAGCTGGCTGGAGGCTTCGTTGAGCTTCGAGGAGCATGGCTGGTAC 555  
Qy 526 TTCGAACTCAGGTTGAAGGCGCTCTCGACCTCTCTCGCTGTTGAGCTTATGGTTCTG 585  
Db 556 ATCGAGATCAGCTCAAGCACACGCGCTGGTGTGGTGGCGCTAGGCTGTC 615  
Qy 586 GAAAGAGATGATCACCGGTGATCACCGGATACCGGAGCCGGGGCGAG 645  
Db 616 GTCGACGGTGGACTACCTGCACTCTGGTGTGGTGGCGCTCTTCAG 675  
Qy 646 GGGCACTAAATCCGCAATGGTTAGGTCTCTGTTGAGGACGACAG 705  
Db 676 GACACCATCAAGTCCGGGGTGGCTGGGGCATAGGAGAC 735  
Qy 706 CTGGCTATCCTGGGGCATCGCTGGTCAAGAGATCAAGTCGTTTCGTTG 765  
Db 736 ATCCGGGTTGCTGTTGGGCCCCGGTGGAGAATCTCAAGTGGCAATCAGGTTTC 795  
Qy 766 AACCTGGTACCTCTGGAGGATCACTGATCGCTGGCGACCTGGTGGCT 825  
Db 796 GAGTCGTTGAACTCTGGCCGCTGGATCTGAGATCTGTCGCGCTGCG 855  
Qy 826 CAGGAATGTTGATGGTAGGTTAACCGCGTGGAGCAACGCCCTGGAGATATC 885  
Db 856 GCGCAACTCGAGCTACACCTGGCAACGAGTAAACGGCGCTGGATGGCTG 915  
Qy 886 ACTCGGATGGAGCTTGTGATTTGCTGGCTGGTGAATGGCCAGTGGCTGGTT 945  
Db 916 CTGGCGTGGGGGGGGTGGAGGGGGTGGTCAATGGTGGGTGAGACGGTGG 975  
Qy 946 TCTACACTCGGGTCACGGCGCACAAGAAAGCGGCCCTATGAGATGGCGTGGC 1005

RESULT 5  
; US-09-718-626-2205  
; Sequence 2205, Application US/09738626  
; Publication No. US20020197665A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAGAWA, SATOHI  
; APPLICANT: MIZOGUCHI, HIROSHI  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIRO  
; APPLICANT: OCHIAI, KEIKO  
; APPLICANT: YOKOI, HARUHIKO  
; APPLICANT: TATEISHI, NAOKO  
; APPLICANT: SEMOH, AKIHIRO  
; APPLICANT: IKEDA, MASATO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-125  
; CURRENT APPLICATION NUMBER: US/09/738,626  
; CURRENT FILING DATE: 2000-12-18  
; PRIORITY APPLICATION NUMBER: JP 99/377484  
; PRIORITY FILING DATE: 1999-12-16  
; PRIORITY APPLICATION NUMBER: JP 00/159162  
; PRIORITY FILING DATE: 2000-04-07  
; PRIORITY APPLICATION NUMBER: JP 00/280988  
; PRIORITY FILING DATE: 2000-08-03  
; NUMBER OF SEQ ID NOS: 7059  
; SOFTWARE: PatentIn ver. 3.0  
; SEQ ID NO 2205  
; LENGTH: 1134  
; TYPE: DNA  
; ORGANISM: Corynebacterium glutamicum  
; US-09-738-626-2205

Query Match 21.4%; Score 239.4; DB 9; Length 1134;  
Best Local Similarity 54.1%; Pred. No. 5.6e-72;保守性 446;  
Matches 533; Conservatve 0; Mismatches 446; Indels 6; Gaps 2;  
Db 22 CAACGTAGAAATCAACAGGTATTAGTTGGATGTCGCGCATGGTCTCC 81  
Db 10 CCACGTGTAACACGCCAACATCGTGGTGGCAAACTGCGCAATCTCA 141  
Qy 82 ATCGCCGTCAGTCATGCCAACATACCGCGTACGACAGAGCTGGCGAAGGCCAACGTC 129  
Db 70 ATTCCGGTCAGTCATGCCAACACCAACGACATACCGGCCCTGCAAC 141  
Qy 142 ATCAAGGGCTGGAACGGTGGCGCTGATACGTCGTTACCGACGATGGC 201  
Db 130 ATCGCACAGTGTGACAGCCCGTGTGACATCGTCCGGCTGGCTGCCAACAGCTG 189  
Qy 202 GCGCAGAGCTTCATCAACCTCATCAAGCAGCTTACGTCGCGCTGGCTGACATC 261  
Db 190 GATGCGAGACACTGCGCATCGCAAGAAGTCTCGATCCGATGGTGGAGATATC 249  
Qy 262 CACTTCGACATTCGATGCGCTGGCAAGATGGCGATTCGGCTGCGATTTGCTG 321  
Db 250 CACTCCGAGCCCTGGTACATCTCGCCCAATCGATGCGTGGCTGGAGATTC 309  
Qy 322 AACCCCTGCAAATC -- GTTAATGAGAGCGCTATCGCAGTGGTGGTGGACTGGCGC 378  
Db 310 ACCGACGAAACATCAAGGAATTCGAGTGGCTGGCTAAAGAGTAGCTGCG 369  
Qy 379 GATAAAACATTGCGTGTGGCTGAAAGTGGCTAACGCCGATCTGGAA -- AGATCTG 435  
Db 370 GATGCCGAAATTCAATTCGATGGCTGGTCAACGGCGATCCCTGGATAACGGCACTCTG 429

QY 436 CAAGAAAGTATGGGACACGACGCCAGCGCAGGGCTGTGGATCTGCCATGGGTCACTGTT 495  
Db 430 GACAATACCAACGCAAACCCACAGAGCTCTGAGTC 489  
QY 496 GATCATCTCGATCGCTGAACCTCGATCACTTCAGAAGTCAGCGTGAAAGGGCTGACGTC 555  
Db 490 GGCCTGTTGAGAGCACGCCCTCGGCACATCGAACCTCTGTGAAAGCAGCTCCGACCA 549  
QY 556 TTCCCTGCTTGTACTCTTCGTTGCGCAAAACACATCGATCAGGGTGCATCG 615  
Db 550 GTACTCATGGGGAGGCCTACCGCAGCTGCTGACAAAGCAGCTCCACTGCACCTC 609  
QY 615 GGGATCACCGAAGGGTGGCGCAGGGGAGTAATCGCCATTTGGTAGT 675  
Db 610 GGTGTACTAACGCTTGTGAGCTGGTCAAGTCAGTGGAAACATGGTCCGAGTC 669  
QY 736 GAAGAGATCAAGTGGTTCTGATATTTCGAAATCGTGTGCTATCCGGTGGGGATC 795  
Db 730 GAAGAAATCAGGTGGCGACAGATTCTCAGTCAGTCCTCACCTCGCCACGGAACTG 789  
QY 676 CTGCTGCTGCTGAAGGCATOGGCACACCCCTGGCGATGGGGCTGTT 735  
Db 670 GCTCAGCTGCCAGGGCATGGCAGTGGGACACTATCCGTGTCTCTGTCGACCGAGTC 729  
QY 796 AACATCATTCCCTGCCAGCTGCTGCTGCCATAGTGGCCATGGTACGGTAAACGCTG 855  
Db 790 GAAATCGTGTCTGCCATAGTGGCCATGGTACGGTGTGACTCTGTA 849  
QY 856 GCGCTGGAGAACGGCTGGAGATATCATCACTCGATGGACGTTTCGATATCGCTGC 915  
Db 850 GAAGTCACCGAACACTCGACGGCATGGGATGGAGTTCACCTGGCGTGTGTCATGGTTC 909  
QY 916 GTGGGAATGGCCCAAGTGTGAGGCGCTGGTCTCAACTCGCGTCAACCGCGGACAACG 975  
Db 910 GTGTGTTAACGCCCAAGTGTGAGGCGCTGGCAGGCTGACCTCGGTGTGCATCCGGTAACGGC 969  
QY 976 AAAAGGGCTCTGAGGTGGCG 1000  
Db 970 AAGGCCAGAGCTTGTCAAGGGCG 994

RESULT 6

US-00-881-752A-195  
; Sequence 195 Application US/09881752A  
; Patent No. US20020115078A1

GENERAL INFORMATION:

APPLICANT: Kleanthous, Harold  
APPLICANT: Al-Garawi, Amal  
APPLICANT: Miller, Charles  
APPLICANT: Tomb, Jean Francois  
APPLICANT: Oomen, Raymond P.

TITLE OF INVENTION: Identification of Polynucleotides  
TITLE OF INVENTION: Encoding No. US20020115078A1 Helicobacter Polypeptides in the  
FILE REFERENCE: 06132-041002  
CURRENT APPLICATION NUMBER: US/09/881,752A  
CURRENT FILING DATE: 2001-06-15

PRIOR APPLICATION NUMBER: US 08/833,457  
PRIOR FILING DATE: 1997-04-01  
NUMBER OF SEQ ID NOS: 370

SOFTWARE: fastSEQ for Windows Version 4.0  
SEQ ID NO: 195  
LENGTH: 190  
TYPE: DNA  
ORGANISM: Helicobacter pylori  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (51)...(1127)

QY 31 AAATCACACGTTAACGTTGGAAATGTCGCCATGGGCTCCATCGCCGTA 90  
Db 69 AGACCAAGGAAATTATCGGGTGGCCATAGGGGTGATCTCCATAGCAGC 128  
QY 91 CAGTCATGACATTCAGCTACAGCAGAGCTGCAAGCAACGGTAACTCACATCGAG 150  
Db 129 CAAACATGACCTTGTGACAAACCGCTGATATTGAAAGCACTAAATCAATTGACAGA 188  
QY 151 CTGGAAGCGTTGGCTGATATCGTCCGTTGATCGTACCGACGGATGGAGCGCAGAA 210  
Db 189 CTCAACTCTGCCGGGGCATGGTTAGTGAGGGTGGGTGAGTAATGAAANGGACCTCTA 248  
QY 211 CGTCAACATCAACATCAACAGCAGCTTAAGCTGCGCTGGCGTGGCGACATCACHTCGAC 270  
Db 249 GCCTTAAGAGATTGAAAAGTGTCCCCTTGCTCTTAATCGCTGATATCTTCAT 308  
QY 271 TATCCGATTCGCTGAAAGTAGCGCAATAGGCGCGATGTCGCGTATAACCTGGC 330  
Db 366 AACATCGGCTTAAGAGAAGTCAAGCGCTGGTGTGAAAGAAAACATT 425  
Db 309 TATAAATTCGCTCTCATGGCGCTCAA--GCGTGGATGCGATCAGGATAACCCCGGA 365  
QY 331 ATATCGGTAATGAAAGAGCTTGGCATGTTGACTGTGCGCGATTAACATT 390  
Db 366 AACATCGGCTTAAGAGAAGTCAAGCGCTGGTGTGAAAGAAAACATT 450  
QY 391 CGCATCGTATTGGCTTAAGCCGGATCGCTGGAAAGAAGTCACTCGAAGAAAGTNTGC 482  
Db 426 OCTATAGAAATTGGCTGAAATGCTGGAGTTAGAAAGCGATTGTCATCRAATA--C 482  
QY 451 GAACCGACGCCGCGCAGGGCTGGCTGCTGAAATCGCCATGGCTGATGTGTGATCATCTGATCGC 510  
Db 483 GGACCCACCCAAAAGCATGGTAAAGAAGCGCTGGTGTGAAAGAAAACATT 542  
QY 511 CTGACTTCGATCACTCAAGTCAGCTCAGCGTGAAGCGCTGACGCTCTCCGCTGTGTGAG 570  
Db 543 TTGGATTTTACCAATTAAAGATTTTAAAGCGAGGCGATGTCGACCATGAA 602  
QY 571 TCTTATGTTGCTGGCAAACAGATCGATCAGCGGTGATCAGGGATCACCGAGCC 630  
Db 603 GCTTACAGGATGCTTGTGCTTGTGATCTATCTTCATTGGGTTACGGGGCC 662  
QY 631 GGTGGGCGCCAGGGCGTAAATCGCCATTGGTTAGGCTGCTGCTGTGAA 690  
Db 663 GGGAAATCTTTTAGCTCCAGTATCACATCGCTATGGCTTTAGGGGCTTTAATGGAG 722  
QY 691 GGCATGGGACACCTGCGTAVGCTGGGGCGATCGGTGAGGATCAAGTC 750  
Db 723 GGCATGGGATACATGCGCTACATCACAGGGAAATTAGAAATGAAATCAAGTG 782  
QY 751 GGTGGATATTGAAATCGCTGGATCGTGGGATCAACTCATCCCTGC 810  
Db 783 GCCAGGCAATTAGCCATAGCGGGCTGAAAGAGGGATPATGGATTCTTCG 842  
QY 811 CGACCTCTGGCTGGCGCTGGATTCAGTGTGTTACCGCGCTGGAGACGC 870  
Db 843 CCCACTTGCGGGCATGACGGCATTTAGTGGATGAGGCTGAAGGTAAACAGC 902  
QY 871 CTGGAGATTCATCACTCGATGGACGTTGATTCGCTGGCTGGGGAATGCCCA 930  
Db 903 TTAAGCCACATCAACCCCTTACATGCGATGGTGGATGGTGAATCGAC 962  
QY 931 GGTGAGGGCTGGTCTACACTCGCGTACCGGGGCAACAGAAACGGCCTCAT 990  
Db 963 GGTGAGGCCAGCATGGCATCCCTTGGATCGCGCTGGGGAATGCCCA 1022  
QY 991 GAAGATGGCGCGAAGACCGCTGGACACAAACGATGATCGA 1037  
Db 1023 AAAGAGGTTAAGTCATCACAACCTGGCTGAAGAGGTTATTGAA 1069

; Patent No. US20020069426A1  
; GENERAL INFORMATION:  
; APPLICANT: Boronat, Albert;  
; APPLICANT: Campos, Narciso;  
; APPLICANT: Rodriguez-Concepcion, Manuel;  
; APPLICANT: Rohner, Michel;  
; APPLICANT: Seeman, Myriam;  
; APPLICANT: Venkatesh, Tyamagondlu V.;  
; APPLICANT: Valentin, Henry E.;  
; APPLICANT: Venkatesh, Mylavarampu  
; TITLE OF INVENTION: Methyl-D-Erythritol Phosphate Pathway Genes  
; FILE REFERENCE: 16516.107/35-21(5189)US  
; CURRENT APPLICATION NUMBER: US/09/921,992  
; CURRENT FILING DATE: 2001-08-06  
; PRIORITY APPLICATION NUMBER: US 60/223,483  
; NUMBER OF SEQ ID NOS: 85  
; SEQ ID NO: 21  
; LENGTH: 584  
; TYPE: DNA  
; ORGANISM: zea mays  
; US-09-921-992-21  
; Query Match 6.3%; Score 70; DB 10; Length 584;  
; Best Local Similarity 52.5%; Pred. No. 9.4e-14;  
; Matches 179; Conservative 0; Mismatches 159; Indels 3; Gaps 1;  
; Qy 595 ATCGATCACCGGTTGCACTGGGATCACCGAAGCCGGTGTGCGAGCGGGACTA 654  
; Db 459 ATGGATTATCCCCTCCACATGGGTACCGGATACTACCGGATA 400  
; Qy 655 AATTCGCCATTGGTTAGGTCTCTGTCAGAACGGATCGGCAGT 714  
; Db 339 TCCCTCACCGGAAGCCCCGAAAGAAATCCGGTTGCTACGCACTCCAGCGCTG 280  
; Qy 715 TCGCTGGCGCCGAMCCGGTTCGAGAGATCAAGTCGGTTCCGATATTGAAATCGCTG 774  
; Db 399 AAATTCACCGCCTCCACATGGGTACCGGATATGGCCGCAAT 340  
; Qy 775 CGTATCCGTTGCGGAGGGATCAACTTCATCGCTGCCGACCTGTCGG 824  
; Db 403 GGTTTGCGGAAACCATGGGAAATATGGGCCCTGCTTCCTGGCCG 452  
; Qy 835 GATGTTATGGTTGGTTAACGCGCTGGGAGAACGCTGGAAAGATATCATCACTCCGATG 894  
; Db 219 AACATTGGAAGACGGTTACATGAGTCCGAGAACCTAAACAT--CTAACGGTTA 163  
; Qy 895 GACGGTTGCAATTATCGGCTGGTGTGATGGCCAGGTGA 935  
; Db 162 GACATCGCCGTCATGGCGTGTCAATTGGCCGGGA 122  
; RESULT 8  
; US-09-921-992-23  
; Sequence 23, Application US/09/921992  
; ; GENERAL INFORMATION:  
; ; APPLICANT: Boronat, Albert;  
; ; APPLICANT: Campos, Narciso;  
; ; APPLICANT: Rodriguez-Concepcion, Manuel;  
; ; APPLICANT: Rohner, Michel;  
; ; APPLICANT: Seeman, Myriam;  
; ; APPLICANT: Valentin, Henry E.;  
; ; APPLICANT: Venkatesh, Tyamagondlu V.;  
; ; APPLICANT: Venkatesh, Mylavarampu  
; ; TITLE OF INVENTION: Methyl-D-Erythritol Phosphate Pathway Genes  
; ; FILE REFERENCE: 16516.107/35-21(5189)US  
; ; CURRENT APPLICATION NUMBER: US/09/921,992  
; ; CURRENT FILING DATE: 2001-08-06  
; ; PRIORITY APPLICATION NUMBER: US 60/223,483  
; ; NUMBER OF SEQ ID NOS: 85  
; ; SEQ ID NO: 22  
; ; LENGTH: 670  
; ; TYPE: DNA  
; ; ORGANISM: zea mays  
; ; US-09-921-992-22  
; Query Match 6.1%; Score 68; DB 10; Length 596;  
; Best Local Similarity 50.1%; Pred. No. 5.5e-13;  
; Matches 129; Conservative 0; Mismatches 101; Indels 0; Gaps 0;  
; Qy 595 ATCGATCACCGGTTGCACTGGGATCACCGAAGCCGGTGTGCGAGCGGGACTA 654  
; Db 223 ATGGATTATCCCCTCCACATGGGTACCGGATACTACCGGATA 342  
; Qy 655 AATTCGCCATTGGTTAGGTCTCTGTCAGAACGGATCGGCAGT 714  
; Db 283 AATTCACCGGTTGCACTGGGATACCGGATTCGATGGGATATCCGGGTA 342  
; Qy 715 TCGCTGGCGCCGAMCCGGTTCGAGAGATCAAGTCGGTTCCGATATTGAAATCGCTG 774  
; Db 343 TCCCTCACCGGAAGCCCCGAAAGAAATCCGGTTGCTACGCACTCCAGCGCTG 402  
; Qy 775 CGTATCCGTTGCGGAGGGATCAACTTCATCGCTGCCGACCTGTCGG 824  
; ; RESULT 9  
; US-09-921-992-22  
; Sequence 22, Application US/09/921992  
; ; GENERAL INFORMATION:  
; ; APPLICANT: Boronat, Albert;  
; ; APPLICANT: Campos, Narciso;  
; ; APPLICANT: Rodriguez-Concepcion, Manuel;  
; ; APPLICANT: Rohner, Michel;  
; ; APPLICANT: Seeman, Myriam;  
; ; APPLICANT: Valentin, Henry E.;  
; ; APPLICANT: Venkatesh, Tyamagondlu V.;  
; ; APPLICANT: Venkatesh, Mylavarampu  
; ; TITLE OF INVENTION: Methyl-D-Erythritol Phosphate Pathway Genes  
; ; FILE REFERENCE: 16516.107/35-21(5189)US  
; ; CURRENT APPLICATION NUMBER: US/09/921,992  
; ; CURRENT FILING DATE: 2001-08-06  
; ; PRIORITY APPLICATION NUMBER: US 60/223,483  
; ; NUMBER OF SEQ ID NOS: 85  
; ; SEQ ID NO: 22  
; ; LENGTH: 670  
; ; TYPE: DNA  
; ; ORGANISM: zea mays  
; ; US-09-921-992-22  
; Query Match 6.1%; Score 58; DB 10; Length 670;  
; Best Local Similarity 56.1%; Pred. No. 5.8e-13;  
; Matches 129; Conservative 0; Mismatches 101; Indels 0; Gaps 0;  
; Qy 595 ATCGATCACCGGTTGCACTGGGATCACCGAAGCCGGTGTGCGAGCGGGACTA 654  
; Db 223 ATGGATTATCCCCTCCACATGGGTACCGGATACTACCGGATA 342  
; Qy 655 AATTCGCCATTGGTTAGGTCTCTGTCAGAACGGATCGGCAGCGCAGCGCTG 714  
; Db 283 AATTCACCGGTTGCACTGGGATACCGGATTCGATGGGATATCCGGGTA 342  
; Qy 715 TCGCTGGCGCCGAMCCGGTTCGAGAGATCAAGTCGGTTGCTGATATTGAAATCGCTG 774  
; Db 343 TCCCTCACCGGAAGCCCCGAAAGAAATCCGGTTGCTACGCACTCCAGCGCTG 402  
; Qy 775 CGTATCCGTTGCGGAGGGATCAACTTCATCGCTGCCGACCTGTCGG 824

Db 403 GGTTGCGGAAACCATGGTGAATATGTCGCCCTGCCTGCGCCG 452

RESULT 10

US-09-974-300-6102

; Sequence 6102, Application US/09974300

; Patent No. US20020146721A1

; GENERAL INFORMATION:

; APPLICANT: Berka, Randy M.

; TITLE OF INVENTION: Methods For Monitoring Multiple Gene

; FILE REFERENCE: 10085\_500-US

; CURRENT APPLICATION NUMBER: US/09/974,300

; CURRENT FILING DATE: 2001-10-05

; PRIORITY APPLICATION NUMBER: 09/680,598

; PRIORITY FILING DATE: 2000-10-06

; PRIORITY APPLICATION NUMBER: 60/279,526

; PRIORITY FILING DATE: 2001-03-27

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO: 6102

; LENGTH: 348

; TYPE: DNA

; ORGANISM: Bacillus clausii

; US-09-974-300-6102

Query Match 5.1%; Score 56.8; DB 10; Length 348;

; Best Local Similarity 54.9%; Pred. No. 4; e-09; Mismatches 92; Indels 0; Gaps 0;

; Matches 112; Conservative 0; Gaps 0;

; QY 764 TCAAATGGCTSGTATCGTGTGGAGGATCACTTCATGCCCTGCCGACCTGTCG 823

; Db 1 TGAATCGTTGCCCTAACGGTAATCGCAGCACACTCATCTCTGCCAACGTGGGG 60

; QY 824 GTCAAGGATTTCATGTTATCGTGTGGAGGATCA 883

; Db 61 GGATAGATGGATTGATGATGATGATGATGAACTGGAAAGTGATATCGAAATAA 120

; QY 884 TCACTGGATGAGCTTTCGATATCGCTGGTGAAGGCCAGGTTGG 943

; Db 121 ANGCACCGGATAAAGTGCTGTCTCGTGTGACTGTAATGCCCAAGGGAGGCCAG 180

; QY 944 TTTCTACACTCCGCGTACCGCCG 967

; Db 181 AGCGAGATCGSTATTCGGAG 204

US-09-921-992-25

; Query Match 4.7%; Score 52.6; DB 10; Length 293;

; Best Local Similarity 53.7%; Pred. No. 1; e-07; Mismatches 94; Indels 0; Gaps 0;

; Matches 109; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

; QY 25 CGTAGAAATCACACGTTAACGTTGGGATGCCCCGATGGGATGGGCTCCCATC 84

; Db 10 CGGAGAAACTCGAAGCTGTGATGGGGATGTCACGTCAGTGAACCCATA 69

; QY 85 GCGTAGCTGATGACCAATACGGCTGACGAGCTGAGAACGCGCAATCAATC 144

; Db 70 AGGATCAACCATGAGACTTCAGATACCAAGGATGTTGGAAACAGTGGAGGGTG 129

; QY 145 AAGGCCTGGAACCGCTTGGCTGTGATATCGTCCGTTGATCGTACCGACATGGACGCG 204

; Db 130 ATGAGGATGCGATAAAGGAGCTGATCTGTTAGATACAGTCAGGTTAGGAGGA 189

; QY 205 GAGAAACGTTCAAACCTACATA 227

; Db 190 GCTGATGCTGCTTGTGAGATCAA 212

RESULT 11

US-09-921-992-25

; Sequence 25, Application US/09921992

; Patent No. US20020069426A1

; GENERAL INFORMATION:

; APPLICANT: Boronat, Albert;

; APPLICANT: Campos, Narciso;

; APPLICANT: Rodriguez-Concepcion, Manuel;

; APPLICANT: Rohmel, Michel;

; APPLICANT: Seeman, Myriam;

; APPLICANT: Valentini, Henry E.;

; APPLICANT: Venkataramesh, Mylavapu V.;

; TITLE OF INVENTION: Methyl-D-Erythritol Phosphate Pathway Genes

; FILE REFERENCE: 16516\_10735-21\_518971US

; CURRENT APPLICATION NUMBER: US/09/921,992

; CURRENT FILING DATE: 2001-08-06

; PRIORITY APPLICATION NUMBER: US 60/223,483

; PRIORITY FILING DATE: 2000-08-07

; NUMBER OF SEQ ID NOS: 85

; SEQ ID NO: 26

; LENGTH: 456

; TYPE: DNA

; ORGANISM: Zea mays

; US-09-921-992-26

Query Match 4.7%; Score 52.6; DB 10; Length 456;

; Best Local Similarity 53.7%; Pred. No. 1; e-07; Mismatches 94; Indels 0; Gaps 0;

; Matches 109; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

; QY 25 CGTAGAAATCACACGTTAACGTTGGGATGCCCCGATGGGATGGGCTCCCATC 84

; Db 10 CGGAGAAACTCGAAGCTGTGATGGGGATGTCACGTCAGTGAACCCATA 69

; QY 85 GCGTAGCTGATGACCAATACGGCTGACGAGCTGAGAACGCGCAATCAATC 144

; Db 70 AGGATCAACCATGAGACTTCAGATACCAAGGATGTTGGAAACAGTGGAGGGTG 129

; QY 145 AAGGCCTGGAACCGCTTGGCTGTGATATCGTCCGTTGATCGTACCGACATGGACGCG 204

; Db 130 ATGAGGATGCGATAAAGGAGCTGATCTGTTAGATACAGTCAGGTTAGGAGGA 189

; QY 205 GAGAAACGTTCAAACCTACATA 227

; Db 190 GCTGATGCTGCTTGTGAGATCAA 212

RESULT 12

US-09-921-992-26

; Sequence 26, Application US/09921992

; Patent No. US20020069426A1

; GENERAL INFORMATION:

; APPLICANT: Albert, Campos;

; APPLICANT: Michel, Rohmel;

; APPLICANT: Myriam, Seeman;

; APPLICANT: Henry E., Valentini;

; APPLICANT: Venkataramesh, Mylavapu V.;

; TITLE OF INVENTION: Methyl-D-Erythritol Phosphate Pathway Genes

; FILE REFERENCE: 16516\_10735-21\_518971US

; CURRENT APPLICATION NUMBER: US/09/921,992

; CURRENT FILING DATE: 2001-08-06

; PRIORITY APPLICATION NUMBER: US 60/223,483

; PRIORITY FILING DATE: 2000-08-07

; NUMBER OF SEQ ID NOS: 85

; SEQ ID NO: 26

; LENGTH: 456

; TYPE: DNA

; ORGANISM: Zea mays

; US-09-921-992-26

Query Match 4.7%; Score 52.6; DB 10; Length 456;

; Best Local Similarity 53.7%; Pred. No. 1; e-07; Mismatches 94; Indels 0; Gaps 0;

; Matches 109; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

; QY 25 CGTAGAAATCACACGTTAACGTTGGGATGCCCCGATGGGATGGGCTCCCATC 84

; Db 10 CGGAGAAACTCGAAGCTGTGATGGGGATGTCACGTCAGTGAACCCATA 69

; QY 85 GCGTAGCTGATGACCAATACGGCTGACGAGCTGAGAACGCGCAATCAATC 144

; Db 70 AGGATCAACCATGAGACTTCAGATACCAAGGATGTTGGAAACAGTGGAGGGTG 129

; QY 145 AAGGCCTGGAACCGCTTGGCTGTGATATCGTCCGTTGATCGTACCGACATGGACGCG 204

; Db 130 ATGAGGATGCGATAAAGGAGCTGATCTGTTAGATACAGTCAGGTTAGGAGGA 189

; QY 205 GAGAAACGTTCAAACCTACATA 227

; Db 190 GCTGATGCTGCTTGTGAGATCAA 212

RESULT 13

US-09-921-992-24

; Sequence 24, Application US/09921992

; Patent No. US20020069426A1

; GENERAL INFORMATION:

; APPLICANT: Albert, Campos;

; APPLICANT: Manuel, Rodriguez-Concepcion;

; APPLICANT: Michael, Rohmel;

; APPLICANT: Myriam, Seeman;

; APPLICANT: Henry E., Valentini;

; APPLICANT: Venkataramesh, Mylavapu V.;

; TITLE OF INVENTION: Methyl-D-Erythritol Phosphate Pathway Genes

; FILE REFERENCE: 16516\_10735-21\_518971US

; CURRENT APPLICATION NUMBER: US/09/921,992

; CURRENT FILING DATE: 2001-08-06

; PRIORITY APPLICATION NUMBER: US 60/223,483

; PRIORITY FILING DATE: 2000-08-07

; NUMBER OF SEQ ID NOS: 85

; SEQ ID NO: 24

; LENGTH: 293

; TYPE: DNA

; ORGANISM: Zea mays

; US-09-921-992-24

Patent No. US20020069426A1  
GENERAL INFORMATION:

APPLICANT: Boronat, Albert;

APPLICANT: Campos, Narciso;

APPLICANT: Rodriguez-Concepcion, Manuel;

APPLICANT: Rohmer, Michael;

APPLICANT: Seeman, Myriam;

APPLICANT: Valentín, Henry E.;

APPLICANT: Venkatesh, Tyamagondlu V.;

APPLICANT: Venkatesh, Mylavavarapu

TITLE OF INVENTION: Methyl-D-Erythritol Phosphate Pathway Genes

FILE REFERENCE: 16516.107/35.21(51897)US

CURRENT APPLICATION NUMBER: 8 US 60/223,483

PRIOR APPLICATION NUMBER: US 60/223,483

PRIOR FILING DATE: 2000-08-07

NUMBER OF SEQ ID NOS: 85

SEQ ID NO 24

LENGTH: 403

TYPE: DNA

ORGANISM: Zea mays

US-09-921-992-24

Query Match 4 4%; Score 48.8; DB 10; Length 564;

Best Local Similarity 51.8%; Pred. No. 3.2e-06;

Matches 172; Conservative 0; Mismatches 142; Indels 18; Gaps 2;

Query 25 GTAGAAATCACCGTATTACCTGGGAAATGCCATTGGGAGTTGGCTCCCATTC 85

Db 131 GGAGAAAACAACAGCATGGATGTTGGTAACGGGTATTTGGTAGGGACATCATAA 190

Query 86 CCGTACAGTCATGACCAATACGGTAGACAGGATGGCTCCCATTC 145

Db 191 GAATCAGGACTGACTGACAGTACAGTACAGTAAAGCATGTTGGAGACATGACAGTTGA 250

Query 146 AGGCCTGGAACCGGTTGGCGNGATATGGCTGGTATCCCTACCGGATGGACCGGG 205

Db 251 TGGATAGCAGATAAAGGAGCTGATATGTACGGATACAGTCAGTCAAGGAGAAGAAG 310

Query 206 CAAAGCGTCAACTCATCAAACAGAGGTT-----ACGGCGGTG 250

Db 311 CTGATGCTGTTGAGATTAACACACCTCTGAGAAAATTACACATACTCTGG 370

Query 311 GTCTGCGTATAACCCGGAAATATCGGTAT 342

Db 428 AGATTCGTTAAACCCCTGGAATTTCGTGAT 459

RESULT 14

US-09-921-992-35

Sequence 35, Application US/09921992

PATENT NO. US20020069426A1

GENERAL INFORMATION:

APPLICANT: Boronat, Albert;

APPLICANT: Campos, Narciso;

APPLICANT: Rodriguez-Concepcion, Manuel;

APPLICANT: Rohmer, Michael;

APPLICANT: Seeman, Myriam;

APPLICANT: Valentín, Henry E.;

APPLICANT: Venkatesh, Tyamagondlu V.;

APPLICANT: Venkatesh, Mylavavarapu

TITLE OF INVENTION: Methyl-D-Erythritol Phosphate Pathway Genes

FILE REFERENCE: 16516.107/35.21(51897)US

CURRENT APPLICATION NUMBER: US/09/921,992

PRIOR APPLICATION NUMBER: US 60/223,483

PRIOR FILING DATE: 2000-08-07

NUMBER OF SEQ ID NOS: 85

SEQ ID NO 16

LENGTH: 379

TYPE: DNA

ORGANISM: Arabidopsis thaliana

US-09-921-992-16

Query Match 4 3%; Score 48.4; DB 10; Length 379;

Best Local Similarity 58.2%; Pred. No. 3.5e-06;

Matches 85; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

Query 598 GATCAGCGGTCATGGGATCACCGACCGGGTGGCGCGAGGGCAATAA 657

Db 215 GATTATCTTCATGGAGTTACTGAGGAGAGGAGATGGCGGATGAAA 274

Query 658 TCCGCCATTGGGTTAGTCTCTGCTGCTGAGGATCGCGACACGGCGGTAAAG 717

Db 275 TCTGGATGAAATGGGACCTCTCAGGACGGGCTCGGTGACACAAAGTTCA 344

Query 718 CTGGGCCATCCGGTCGAGAGAT 743

Db 335 CTGACGGGCCACAGAAGAGGAGAT 360

US-09-921-992-35

Mon May 12 08:44:42 2003

us-09-921-992-3.rnpb

Search completed: May 10, 2003, 17:56:02  
Job time : 630 secs

